Structures 75 (2025) 108826

Contents lists available at ScienceDirect

Structures

journal homepage: www.elsevier.com/locate/structures

Machine learning-based rapid damage states assessment and seismic response prediction of highway bridges supported by unbonded laminated rubber bearing

Ranxing Gu [a], Junfeng Jia [a,*], Jiachen Bian [a], Tong Guo [b]

[a] State Key Laboratory of Bridge Safety and Resilience, Beijing University of Technology, Beijing 100124, China
[b] Key Laboratory of C & PC Structures of Ministry of Education, Southeast University, Nanjing 214135, China

ARTICLE INFO

Keywords:
Machine learning
Ensemble learning
Deep neural network
TabNet
Seismic isolation bridges
Seismic response prediction
Seismic damage states assessment

ABSTRACT

Highway bridges supported by the unbonded laminated rubber bearings (ULRBs) are critical components of the transportation infrastructure systems. Accurate and efficient assessment of damage states and seismic response prediction for individual bridges at regional scale is essential for evaluating seismic vulnerability and enhancing infrastructure resilience. However, such assessments are typically resource-intensive, requiring significant labor, time, and computational power. This study introduces an approach based on machine learning (ML) algorithms to develop predictive models for rapid damage states assessment and seismic response prediction, incorporating various types of input data. The ML includes five ensemble learning algorithms and three deep neural networks (DNNs), notably TabNet, a novel architecture optimized for tabular data processing. Using a comprehensive dataset generated through nonlinear time-history analyses (NTHA) of 120 ULRB-supported highway bridges subjected to 320 ground motions, a total of 38,400 data samples were obtained for model training and evaluation. Results demonstrate that TabNet outperformed other models in predicting damage states, achieving accuracies of 93.6 % and 90.5 % across two test sets. For predicting the displacement of ULRBs, CatBoost demonstrated superior performance, achieving R^2 values of 0.949 and 0.905 for the two test sets. Furthermore, model interpretability was enhanced using SHAP analysis, which identified the ULRB friction coefficient, elastic stiffness, and spectral accelerations at 0.8 s and 1.0 s as key predictors for bridge damage states and seismic response. This study represents the first attempt to apply TabNet for damage assessment and response prediction in highway bridges. The results provide a valuable reference for applying ML techniques to evaluation of highway bridge performance under earthquake events.

1. Introduction

Earthquakes are among the most destructive natural disasters, causing significant human casualties and economic losses over recent decades [1]. Accurate assessment of damage states and prediction of seismic responses for bridges are critical to mitigating these losses, relying heavily on precise and timely predictions [2]. However, due to the uncertainties of dimensional size, materials, configuration, and ground motion (GM), the accurate prediction of damage states and response of structure in a given region faces great challenges [3–5]. Traditionally, damage states and seismic responses are evaluated using simplified methods, such as the capacity spectrum method (CSM) outlined in codes [6], or through nonlinear time history analysis (NTHA). While CSM is straightforward and efficient, its reliance on a single seismic intensity measure (IM) can lead to deviations of up to 61 % in vulnerability curves [7]. Although NTHA offers higher accuracy, it is impractical for emergency response scenarios due to its complex modeling requirements and substantial computational demands. Furthermore, acquiring the numerous parameters necessary for NTHA modeling is often challenging in real-world engineering applications. Thus, achieving a balance among prediction accuracy, computational efficiency, and implementation complexity is essential.

The advent of artificial intelligence (AI) has profoundly impacted industries and society [8]. Within AI, machine learning (ML) has gained increasing attention in civil engineering due to its exceptional ability to address nonlinear problems [9–11]. ML methods have been successfully

* Corresponding author.
E-mail addresses: guranxing@emails.bjut.edu.cn (R. Gu), jiajunfeng@bjut.edu.cn (J. Jia).

https://doi.org/10.1016/j.istruc.2025.108826
Received 23 December 2024; Received in revised form 27 March 2025; Accepted 29 March 2025
Available online 7 April 2025
2352-0124/© 2025 Institution of Structural Engineers. Published by Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

applied in seismic signal processing [12–14], structural health monitoring [15,16], collapse state assessment [17,18], and seismic response prediction for buildings and bridges [19–22]. These applications demonstrate the potential of ML to address significant prediction biases caused by limited IMs [18–21].

Tree-based ensemble learning methods, known for their efficiency, interpretability, robustness, and stability, have been widely adopted in fields such as finance [23] and healthcare [24]. It typically demonstrates excellent performance in processing tabular data such as IMs [25]. Recent advancements have established ensemble learning as a preferred data-driven analysis technique for assessing damage states and predicting seismic responses. Chen et al. [26] used Natural Gradient Boosting (NGBoost) to predict the strength of low-rise shear walls and classify the seismic damage states in bridges. The results showed that the prediction accuracy of NGBoost is comparable to that of traditional ML models, while providing robust estimates for uncertainty in prediction. Nguyen et al. [27] used and compared eight ML models, including five tree-based ensemble methods, for a cluster of steel-framed structures. The results showed that XGBoost achieved the best performance with a 97 % accuracy rate. Similarly, Xu et al. [28] demonstrated the effectiveness of ensemble learning, particularly XGBoost, in regional seismic performance assessment of reinforced concrete (RC) bridges. Zhang et al. [29] employed active learning and ensemble models for rapid damage states assessment of RC frames, with XGBoost and active learning both performing exceptionally well. Despite these advantages, ensemble learning often requires complex feature engineering, relying on prior knowledge to extract tabular features from GM records.

Deep neural networks (DNNs) address this limitation by enabling automatic feature extraction from GM records without prior knowledge. Convolutional neural networks (CNNs), a prominent DNN architecture, have demonstrated superior learning and generalization capabilities in image-based damage detection [30,31]. Wu et al. [32] used a deep CNN architecture to predict the response curve of the single degree of freedom (SDOF) system. Compared to the multilayer perceptron (MLP), CNN exhibited less error and greater robustness. In addition, a physical-level explanation was provided. The authors demonstrated that the integral operator can be approximated by a convolutional layer, and the convolutional kernel is more robust as it eliminates more irrelevant information and retains the main frequency features of the entire network. Oh et al. [33] proposed a CNN-based prediction model for seismic response estimation, while Mangalathu et al. [34] utilized wavelet-transformed GM data to predict damage states in buildings and bridges. However, separate CNNs for each structure may be inefficient when analyzing structural clusters. Wen et al. [35] proposed a CNN-based prediction framework for RC frame clusters that establishes a mapping between GM and structural parameters. This prediction model extracts features from GM records through multiple convolutional layers, then concatenates the structural features, all of which are input into a fully connected (FC) layer. The most advantage of this approach is that it takes into account structural uncertainties along with uncertainty of GM, and can capture changes in structural parameters and GM records simultaneously. Zhang et al. [36] used a physics-guided CNN for response prediction (phyCNN), where embedded physical information (e.g., dynamical control laws) can provide guidance for learning and constrain the training to a feasible space. In the test on two different SDOF systems with the 6-story hotel building in San Bernardino, California, phyCNN demonstrated good reliability and robustness. One solution to the limitation of CNNs terms of global information is to add recurrent neural networks (RNN) [37]. Nie et al. [38] combined CNN with a RNN layer to develop a classification and identification method for a real-time seismic signal monitoring system. The study demonstrated that the CNN-DBiLSTM model exhibited the best performance. Zhang et al. [39] introduced a hybrid prediction model combining CNN and long short-term memory (LSTM) to predict the seismic response of high-speed railways. The CNN-LSTM was compared with the Backpropagation (BP) neural network, LSTM, and gated recurrent unit (GRU) in the study, and it was found that RNNs may be more advantageous after the introduction of CNNs in feature extraction. CNNs are not only able to extract features from unstructured data and achieve end-to-end learning but also have strong scalability and flexibility.

Although ensemble learning methods and DNNs have demonstrated significant potential, no existing studies have systematically compared their performance in rapid damage states assessment and seismic response prediction for bridge clusters. This study addresses this gap by proposing a universal prediction framework applicable to both ensemble learning and DNNs. Notably, this study introduces TabNet, a novel DNN architecture specifically optimized for tabular data, which combines the accuracy and efficiency of ensemble learning with the scalability and flexibility of DNNs. The main contributions of this study are as follows:

1. A data-driven framework is proposed for rapid damage states assessment and seismic response prediction. The framework integrates advanced ML algorithms, including tree-based ensemble methods (e.g., Random Forest (RF) and Extreme Gradient Boosting (XGBoost)) and DNNs (e.g., TabNet and ResNet-MLP).
2. TabNet, a novel DNNs architecture, is introduced in bridge seismic performance evaluation. TabNet demonstrates superior performance in damage states assessment, combining the interpretability and efficiency of ensemble models with the flexibility and scalability of DNN architectures.

2. Methodology

Fig. 1 presents the proposed framework for post-earthquake bridge damage identification and seismic response prediction. This framework is designed to rapidly forecast the damage states and seismic responses of structural groups with statistical significance. The process involves the following stages:

1. **Seismic Source and Data Collection:** The framework begins by identifying potential seismic source zones and corresponding seismic intensity levels for the target engineering region. Subsequently, seismic data and relevant structural information for the target structures within the defined area—such as medium and short-span simply supported beam bridges or frame structures—are collected.
2. **Structural Modeling and Dataset Construction:** Structural models and seismic records are randomly paired to conduct nonlinear time-history analysis (NTHA), enabling the extraction of seismic responses and corresponding structural damage states. These outputs are used to construct a structural seismic response dataset. Structural rapid damage assessment is modeled as a classification problem, while seismic response prediction is treated as a regression problem.
3. **Training and Testing Predictive Models:** ML models are trained to identify bridge damage states and predict seismic responses. The predictive modeling leverages structural and seismic information as input, using the dataset constructed in Step 2. Cross-validation is employed to select the best-performing model.
4. **Practical Application of Predictive Models:** The trained predictive models serve two primary purposes: pre-earthquake assessments and post-earthquake evaluations. Before an earthquake, the models can identify potential vulnerabilities in bridge structures within specific regions, enhancing seismic resilience and mitigating potential damage. After an earthquake, the models facilitate rapid assessment of damage states across bridge networks, supporting efficient recovery of transportation systems and aiding societal recovery efforts.

It is important to note that while the focus of this study is on bridges, the same approach can be applied to other statistically significant structural forms, such as steel moment frames. The framework facilitates the introduction of new instances, thereby enhancing the model's performance. It mitigates the risk of overfitting and establishes a self-

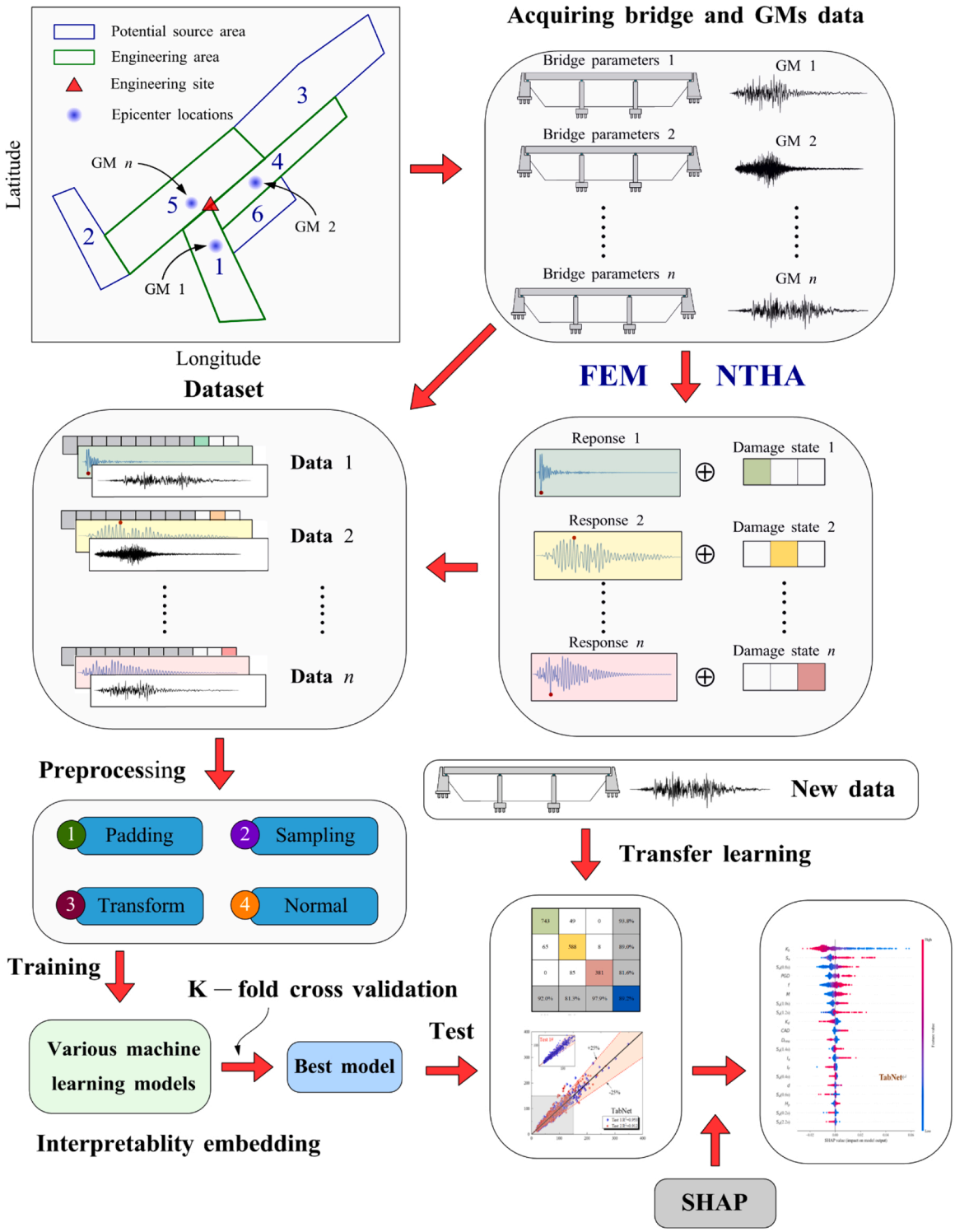


Fig. 1. Workflow of the proposed method.

regulating system that is continuously optimized.

3. Dataset construction

3.1. Description of the ULRBs supported highway bridges

The ULRBs supported highway bridges , particularly small-and-medium-span continuous girder bridges, are widely used in China [40, 41]. They are also found in parts of the central and eastern United States [42]. As shown in Fig. 2 **(a)**, laminated rubber bearings (LRBs) are typically placed directly between the superstructure and substructure, with no restraints other than friction. This configuration enables the dissipation of earthquake energy through bearing sliding, providing seismic isolation for the bridge [40,41].

However, the sliding behavior of LRBs increases the displacement response of the superstructure, elevating the risk of unseating [43,44]. To mitigate this, concrete restraining devices are commonly employed as reinforcement [45]. Nevertheless, the relative brittleness and limited

Fig. 2. (a) Sliding of laminated-rubber bearings [40]; (b) Reinforcement by TYSD.

energy dissipation capacity of concrete restraining devices can increase substructure response demands during large earthquakes, potentially leading to damage that is difficult to repair [46].

Triangular yielding steel dampers (TYSDs), which offer high energy dissipation capacity and are relatively easy to replace, have been demonstrated as an effective alternative for transverse restraining devices in bridges [41,47,48]. As shown in Fig. 2**(b)**, incorporating TYSDs in ULRBs-supported highway bridges is a feasible solution for enhancing transverse reinforcement and mitigating damage risks.

3.2. FE Simulation of bridge

ULRBs-supported highway bridge typically employs a three-span continuous girder with six lanes in both directions and most of them are designed to be of the same span [45]. For this study, 3D finite element (FE) models were developed using the OpenSeesPy platform (a Python-based third-party library for OpenSEES), as shown in Fig. 3. In the seismic isolation system comprising a continuous girder bridge with ULRBs and TYSDs, the girders and bent caps were assumed to remain elastic during an earthquake. Therefore, Elastic Beam-Column Elements were used for the simulation of girders and bent caps [22,49–51]. The geometrical parameters of the box girder were calculated in different directions and assigned, along with the mass, to the longitudinal elements at the geometrical center of the deck. Rigid Links were used to simulate the box girder diaphragm to ensure the transfer of moments and forces, and to rigidly connect them with the superstructure [45,52].

Since the damage to bents in seismic isolation bridges is typically minimal, it is reasonable to use Elastic Beam-Column Elements for simulation purposes [45]. However, the additional TYSDs increases the displacement response of bents under strong earthquake, thereby increasing the possibility of columns reaching a plastic status. Therefore, in this study, Fiber-Type Displacement-Based Beam-Column Element was employed to simulate the bridge column to consider its non-linear behavior [41]. Where concrete was defined using the Concrete 01, while steel reinforcement was defined using Steel02 [50].

In this study, the bents were supported by two columns, which were in turn supported by a group of piles with pin-type connections at the foundation. These connections were simulated by linear horizontal and rotary springs placed at the center of the foundation and the torsional stiffness was set to 0, in order to describe the behavior of such structures [3,53]. The bridge abutments exhibit different mechanical behaviors in the longitudinal and transverse directions. In the longitudinal direction, the pressure of backfill soil and the piles act together in the passive behavior, whereas only the piles carried the load in the active behavior. The soil contribution in the transverse direction was also neglected, allowing the pile to move in this direction [21]. The behavior of the pile was modeled using zero-length elements with the trilinear model proposed by Choi et al. [53]. The backfill behavior of the soil was modelled using zero-length elements, based on the hyperbolic gap model proposed by Shamsabadi and Yan et al. [54]. Furthermore, the girder will exhibit a significant displacement under strong earthquakes which will have an impact on the abutment. Therefore, the Hertz collision model proposed by Muthukumar and DesRoches [55] was employed to simulate the collision between the girder and the abutment, using impact material elements.

In each abutment and bent, six LRBs were arranged, with one TYSD on each side of them. Since LRBs exhibited sliding behavior under the strong earthquake, they are subjected to simulation using flatSliderBearing elements. Where the transverse and longitudinal mechanical behaviors were set as linear elastic spring, and the sliding behavior of the LRBs was described by the Coulomb friction model. TYSDs were simulated using zero-length elements. Its transverse mechanical behavior was represented by a bilinear model implemented in the Steel01. One end node of these element was connected to the pier-cap node, and another node was connected to the superstructure node by a rigid link element [41,45].

The reasonableness of this modelling approach has been verified by a multitude of studies [3,20,21,45,50,51,53,56]. More detailed modeling information can be found in the studies by Mangalathu [3], Choi [53], and Ramanathan [56].

3.3. Modeling the uncertainties

To develop a reliable prediction model, it is essential to account for the uncertainties associated with earthquakes, materials, and structural properties. For ULRBs-supported small-and-medium-span continuous beam bridges in China and the central and eastern United States, relevant literature was extensively reviewed. Twenty-two easily accessible geometric and material variables were identified, and their statistical distributions were established, as summarized in Table 1.

In order to consider the GMs uncertainties, 320 GM records were selected from the database developed by Baker et al. [62]. The dataset comprises four sets of 80 records each (namely 1 A, 1B, 2, and 3). It considers the effects of seismic intensity, soil type, magnitude, focal depth, and fault type (normal and parallel). Notably, near-fault GMs characterized by pulse-like features have significant effects on structural seismic responses. To consider the influence of near-fault effects, 80 pulse-like GM records (Set #3) were deliberately included. Fig. 4 shows the 5 % damped acceleration response spectrum for 320 GM records. More detailed information is available in Baker et al. [62]. The acceleration response spectra of these selected GMs, including pulse-like records, cover a broad spectral range and demonstrate substantial variability in amplitude and frequency content. This ensures that the selected dataset adequately captures the complex characteristics of both near-fault and far-field GMs. To ensure that the displacement response decays to a stable value after earthquake, all GM records were filled with 30 seconds of zeros at the end. The additional segment was consistent with the time step of the original records.

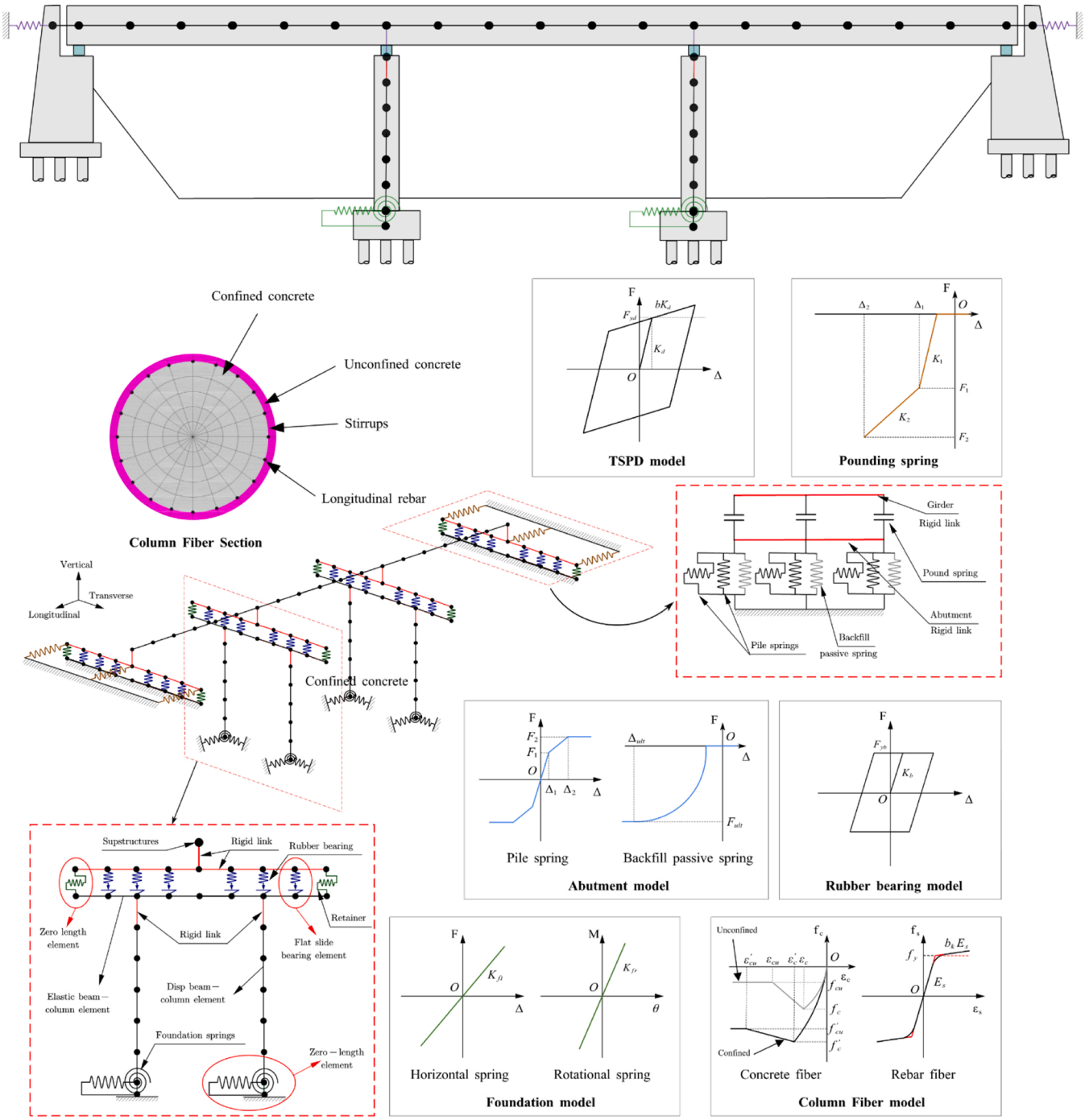


Fig. 3. FE model of ULRBs-supported continuous girder bridge with TYSD.

3.4. Generating realistic bridge specimens

The 3D bridge specimens were generated using Latin Hypercube Sampling (LHS), a statistical method providing more evenly distributed samples across the entire parameter space than traditional Monte Carlo sampling [57]. Specifically, the Python package pyDOE was employed to implement LHS. A total of 22 structural parameters listed in Table 1 were sampled. The cumulative distribution function (CDF) of each parameter was divided into 120 equal segments. A single random value is generated within each of the segments. Subsequently, the inverse CDF is applied to transform the uniformly distributed value into the corresponding value in the target distribution. To ensure reproducibility, the random seed was fixed at 42 (random_state=42). Subsequently, samples outside the predefined 95 % confidence intervals of the probability distributions were rejected. Specifically, the acceptance criterion was that all generated values must fall within the lower and upper bounds defined by the cumulative probability limits corresponding to the 2.5th and 97.5th percentiles of the respective distribution of each parameter.

A total of 120 realistic bridge specimens were generated using LHS, with each specimen modeled following the approach outlined in Section 3.1. Consequently, the dataset comprised 120 bridge specimens paired with 320 GM records, resulting in a total of 38,400 samples (120 × 320). For the NTHA performed in OpenSeesPy, the time step was set consistently equal to the sampling interval of the GM records. Rayleigh damping with a 5 % damping ratio at fundamental modal frequency was uniformly applied to all bridge specimens. The integration scheme

Table 1
Statistic distribution of structural and material variables.

		Description	Probability distribution	Mean	Cov	Lower	Upper	Reference
Supstructure	M	Superstructure mass (ton)	Uniform	1402.5	10.00 %	1170	1635	[56]
	L_s	Span length (m)	Uniform	30	10.00 %	25	35	[56]
	w_g	girder width (m)	Uniform	17.37	14.05 %	15.24	20.01	[57]
Column	H_p	Column height (m)	Normal	8.8	32.00 %	1.62	16.25	[57]
	d_p	Column diameter (m)	Normal	2	10.00 %	1.46	2.61	[58]
	ρ_L	Longitudinal reinforcement ratio of column	Normal	0.015	27.00 %	0.0007	0.0284	[59]
	ρ_T	Transverse reinforcement ratio of column	Normal	0.005	42.00 %	0.0002	0.0105	[59]
Rubber bearing	K_b	Initial stiffness of a plate bearing (kN/mm)	Uniform	2.95	41.00 %	0.85	5.05	[57]
	f	Friction coefficient	Uniform	0.2	29.00 %	0.1	0.3	[57]
transverse steel dampers	K_d	Initial stiffness of a TYSD (kN/mm)	Truncated Lognormal	4.07	71.00 %	0.525	19.3	[57]
	F_{yd}	Equivalent yield strength of a TYSD (kN)	Truncated Lognormal	49	60.00 %	10	180	[57]
	b	Post-yield stiffness ratio	Lognormal	0.08	39.00 %	0.027	0.221	[57]
Material	f_c	Concrete strength (MPa)	Lognormal	34.5	18.00 %	19.92	58.05	[60]
	f_y	Reinforcement strength (MPa)	Lognormal	420	10.60 %	319.23	549.14	[61]
	E_s	Elastic modulus of rein. (MPa)	Lognormal	201	3.30 %	184.28	219.6	[61]
	b_k	The hardening ratio of reinforcement	Lognormal	0.0083	43.00 %	0.0026	0.0236	[56]
Abutment & pile	K_{ft}	Foundation translational stiffness (kN/mm)	Lognormal	175.1	0.25 %	70.05	437.8	[21]
	K_{fr}	Foundation rotational stiffness (GN-m/rad)	Lognormal	1.36	20.59 %	0.54	3.39	[21]
	K_{pi}	Piles translational stiffness (kN/mm/pile)	Lognormal	0.124	36.29 %	0.058	0.234	[21]
	K_a	Abutment passive initial stiffness (kN/mm/m)	Uniform	21.75	16.09 %	14.5	29	[20]
	H_a	Abutment backwall height (m)	Lognormal	3.59	18.11 %	2.9	6.1	[20]
	g_L	Longitudinal gap (mm)	Lognormal	23.3	53.22 %	7.6	55.9	[20]

employed in NTHA was Newmark's average acceleration method.

4. Preparation of predictive models

4.1. Model architectures and data fusion methods

To accurately predict the damage states and seismic responses of a structural portfolio, it is crucial to incorporate both structural characteristics and GM information into the ML model. Structural parameters are typically expressed as tabular data, whereas GM data can be represented either by engineered features extracted based on prior knowledge or by raw time-series data processed through DNNs. As illustrated in Fig. 5, two principal methods were employed to fuse structural and seismic information: 1. Tabular-based feature fusion: GM features (e.g., *PGA*, S_a, *EDP*, etc.) are extracted using established engineering practices and merged with structural characteristics and site categories. These fused datasets are then processed by models that are adept at processing tabular data, such as tree-based ensemble algorithms. 2. Direct time-series feature extraction using DNNs: Raw GM records are directly input into DNNs (e.g., CNN, ResNet) for automatic feature extraction. The extracted high-dimensional features are then connected to a decoder, typically MLP, for output.

4.2. Input and output variables definition

In the context of the general DNN architecture, the input data is divided into three distinct categories: tabular data (bridges structural features, e.g., span length, column height, superstructure mass), categorical data (site class, fault type), and time series data (GM records). The bridge structural features are min-max normalized before input, which scales the different feature value ranges to similar intervals, thereby improving the model's performance. Categorical feature data is processed using one-hot encoding, a technique that transforms categorical data into a numerical format. This type of feature extraction, based on the "bag of words" model, addresses the challenge of discrete and unordered categorical features. Given the wide range of GM

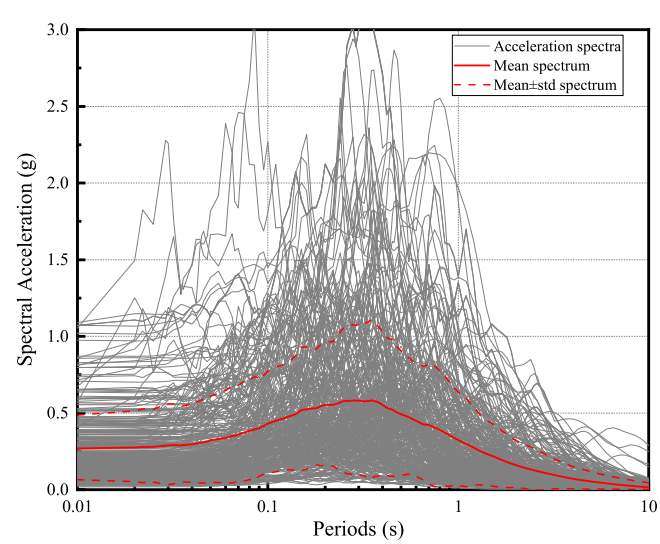


Fig. 4. Acceleration respond spectra.

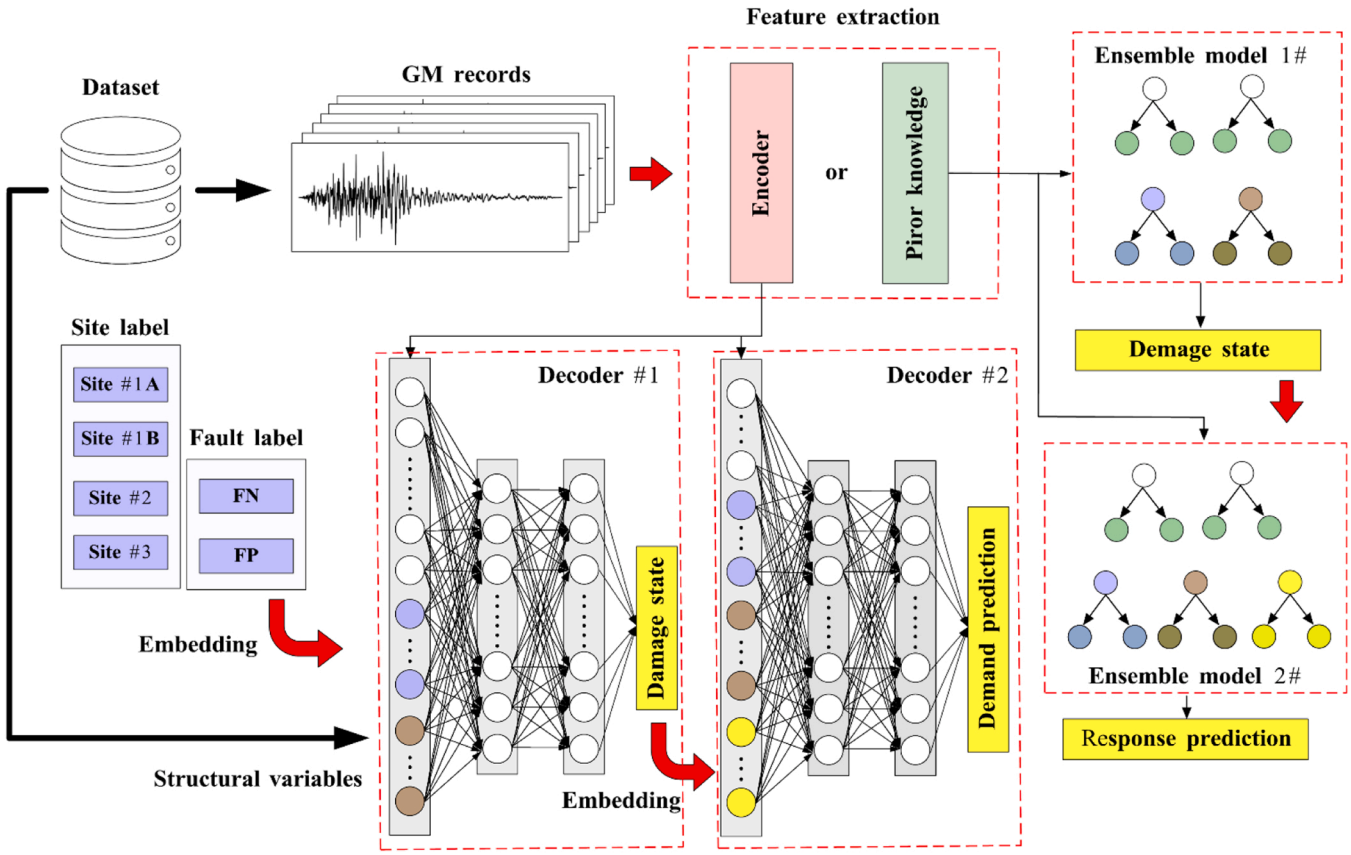


Fig. 5. Model architectures and two data fusion methods.

sampling frequencies, calculating appropriate model weights can be challenging [28]. To facilitate ML operations, it is necessary to unify the sampling frequencies and record duration. Relevant studies have shown that GM records with a sampling frequency of 50 Hz and a duration of 30 seconds are typically sufficient for most earthquake engineering applications [63]. Therefore, a down-sampling frequency of 50 Hz was applied to the original GM records, resulting in input data with dimensions of (1500, 1).

For the ML model processing tabular data, structural and categorical features were processed similarly. A total of 45 GM intensity measures were selected, including peak ground acceleration (*PGA*), effective peak acceleration (*EPA*), spectral intensity (*SI*), and values of the acceleration response spectrum with a damping ratio of 5 % at period 0.2 s~3.0 s ($S_a(0.2s)$~$S_a(3.0s)$). These GM features were considered as the important intensity measures in the past research [17,19]. The specific input and output variables are presented in Table 2.

Table 2
Input and output variables of the ML model.

Variable	Notation		
Input	Ground motion features	Tabular data	AI, A_{rms}, A_{rs}, A_{sq}, CAA, CAD, CAV, D_{rms}, D_{rs}, D_{sq}, EPA, I_a, I_c, I_{ch}, I_d, I_F, I_v, I_z, P_a, P_d, PGA, PGD, P_v, $Ratio$, S_a, $S_a(t)(t = [0.2: 0.2: 3.0]s)$, d, V_{rms}, V_{rs}, V_{sq}
		Sequence data	GM records ([0.05: 0.05: 3000])
		Categorical data	Earthquake sets and fault types
	Bridge features	Tabular data	M, L_s, w_g, H_p, d_p, ρ_L, ρ_T, K_b, f, K_d, F_{yd}, b, f_c, f_y, E_s, b_k, K_{ft}, K_{fr}, K_{pi}, K_a, H_a, g_L
Output	Regression	Tabular data	Peak displacement of bearing sliding ($D_{b,\max}$)
	Classification	Categorical data	No damage (ND)、bearing sliding (BS)、bearing failure (BF)

4.3. Models for time series data

CNNs are effective at capturing complex and abstract features [64]. The characteristics of convolution operation to achieve local receptive fields and weight sharing significantly reduce the number of parameters in the model, thereby decreasing computational complexity and reducing the risk of overfitting [65]. It is therefore common practice to use CNNs, or variants of CNNs, as an encoder. This study refers to the network architecture from Wen et al. [35], utilizing a four-layer CNN as the encoder to process GM records. The structural information is then fused and an MLP is used as the decoder for output.

CNNs, while demonstrating strong performance across a wide range of complex tasks, have certain limitations, including the risk of vanishing or exploding gradients as network depth increases [66,67]. To mitigate these challenges, the introduction of residual connections in CNNs has been proposed as an effective solution [68]. Residual networks (ResNets) enable the training of deeper networks, facilitating easier training and improving the model's expressive power [69–72]. Yun et al. [15] demonstrated the efficacy of combining ResNet with MLP for multi-dimensional signal fusion tasks.

In this study, we enhanced the encoder component of the network architecture proposed by Wen et al. [35] by integrating the ResNet-MLP architecture, as shown in Fig. 6. Specifically, 1D CNNs were used for GM time-series data. The initial convolutional layer employs 1 × 5 kernels with a stride of 2, followed by a max-pooling layer with kernel size of 3 and a stride of 2. Subsequently, the encoder comprises four Res-Blocks. With each Res-Block, the feature length was halved and the number of channels doubled, enhancing feature extraction capability. After the

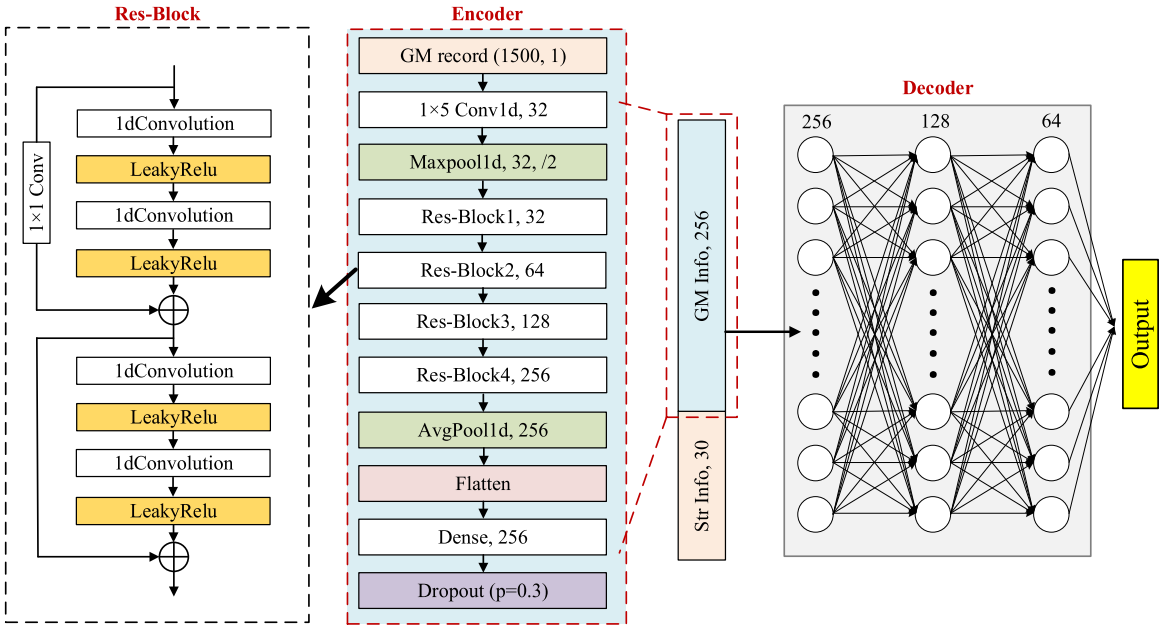


Fig. 6. The architecture of the ResNet-MLP.

Res-Blocks, a global average pooling layer reduced the feature map dimensions to 1, followed by a flattening layer that passes the features into a dense layer. To improve generalization for unknow GM records, a dropout layer was employed. Finally, the GM features extracted by encoder are concatenated with the structural information and then processed through an MLP decoder with three FC layers. The application of nonlinear activation functions enables the network to flexibly learn the complex mappings between inputs and outputs [73]. LeakyRelu activation functions are adopted to effectively mitigate neuron saturation and gradient vanishing issues. It is worth noting that Batch Normalization (BN) did not yield satisfactory results in practice, as mentioned in the study by Wen et al. [35]. Therefore, BN was not used in this model.

Additionally, modifications had been made to the network proposed by Wen et al. [35]. Overall, the CNN-MLP and ResNet-MLP architecture were identical in form, with the only difference being that the Res-Block in the encoder, as shown in Fig. 6, were replaced by CNN layers with the same channels.

4.4. TabNet: a novel DNN model for tabular data

Despite the significant successes of deep learning in various domains, including the processing of time series data such as GM records, it has not been as effective for handling tabular data [74]. Tree-based ensemble models, including Random Forest (RF), Adaptive Boosting (AdaBoost), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting (LightBoost), and Categorical Boosting (CatBoost), have traditionally been favored for tabular data because of their interpretability, robustness, and superior predictive performance compared to conventional DNNs [75–77]. However, ensemble models rely heavily on manual feature engineering and domain expertise, which limits their flexibility and applicability in complex scenarios requiring automated feature learning or domain adaptation [78].

To overcome these limitations, this study introduces TabNet, a novel deep neural network architecture specifically designed for tabular data, leveraging a sparse attention mechanism for adaptive feature selection and end-to-end learning [78], as shown in Fig. 7. Specifically, the original features are processed through BN layer, after which the same dimensional features are passed to each decision step. The encoder of TabNet is comprised of three components: a feature transformer, an attentive transformer, and feature masking. These components operate at each step i of the encoder. The feature representation processed by the feature transformer is divided into attentive transformer inputs for subsequent steps and constitutes a portion of the overall outputs. The decoder of TabNet is constituted by a feature transformer and a FC layer at each step j.

TabNet uses a learnable mask for soft selection of the salient features, the learning capacity of a decision step is not wasted on irrelevant features, and thus the model becomes more parameter efficient. Mask is realized by the multiplication $M[i]\cdot f$. In the network, the masks are obtained by using the processed features from attentive transformer in the preceding step:

$$M[i] = \text{sparse max}(P[i-1]\cdot h_i(a[i-1])) \tag{1}$$

The Sparsemax normalization method encourages sparsity by mapping the Euclidean projection onto the probabilistic simplex. This approach has been observed to perform better than other methods and is aligned with the goal of sparse feature selection for explainability. h_i is a trainable function represented by a FC layer and a BN layer. $P[i]$ is the prior scale term, denoting how much a particular feature:

$$P[i] = \prod_{k=1}^{i}(\gamma - M[k]) \tag{2}$$

Where γ is a relaxation parameter. As γ increases, more flexibility is provided to use a feature at multiple decision steps.

TabNet uses feature transformer to process the filtered features and then splits them into decision step output $d[i]$ and next step input $a[i]$:

$$[d[i], a[i]] = f_i(M[i]\cdot f) \tag{3}$$

In order to achieve parameter-efficient and robust learning with high capacity, a feature transformer should incorporate layers that are shared across all decision steps. Furthermore, it should incorporate layers that are specific to each decision step in order to facilitate step-dependent processing. A feature transformer comprises two shared layers and two decision step-dependent layers. Each layer includes an FC layer, a BN layer, and a GLU, which are ultimately connected to a residual connection with normalization. Normalization with a value of $\sqrt{0.5}$ serves to stabilize learning by ensuring consistent gradient updates. TabNet permits the selection of features through the utilization of sparse instances, which have been learned from the data. Furthermore, it

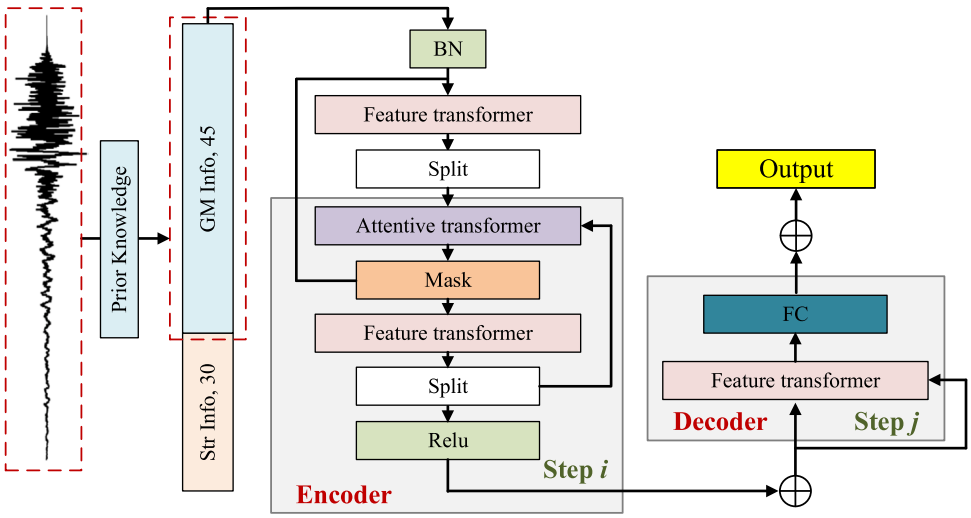


Fig. 7. The architecture of the TabNet.

facilitates enhanced learning through the non-linear processing of the selected features and enables the simulation of ensemble algorithms through higher dimensions and greater number of steps. Further details regarding TabNet can be found in the study by Arik et al. [78].

4.5. Model training and hyperparameter optimization

To robustly evaluate the predictive performance and generalization capability, two test sets were designed. To evaluate the predictive generalization for unseen seismic events, a total of 1920 groups were created by assigning 5 % of the ground shaking records from each site and fault type to Test Set 2#. These records were not used in training or validation. The remaining data were partitioned into training (80 %), validation (10 %), and an additional test set (Test Set 1#, 10 %). This explicit configuration ensures that model performance is rigorously evaluated under both familiar and novel conditions.

Based on the input and output variables listed in Table 2, open-source libraries such as Scikit-learn [79] and Pytorch [80] were used to implement various ML models.

One crucial hyperparameter in the training process is the learning rate (LR). A high LR can lead to unstable training and model divergence, while a low LR may result in slow convergence and the risk of getting stuck in local minima. To address this, the Adam optimizer, a widely-used adaptive learning rate optimization algorithm, was employed [81]. Adam combines the advantages of Momentum and Root Mean Square Propagation, adjusting the LR for each parameter based on the mean of the squared historical gradients. This adaptive nature of Adam helps ensure stable and efficient training, making it well-suited for DNN optimization. A learning rate decay strategy was also applied to improve the convergence of the model. The initial LR and the decay rate were treated as hyperparameters for optimization. In addition, an early stopping strategy was implemented to prevent overfitting.

In this study, Optuna, an efficient and versatile hyperparameter optimization framework [82], was adopted to systematically explore optimal hyperparameters for all models. The detailed hyperparameters and their search spaces for all models optimized using Optuna are summarized in Appendix Table A. Optimization was guided by objective metrics: cross-entropy loss was minimized for classification tasks, while Huber loss was minimized for regression tasks. Each optimization run consisted of 100 trials per model, leveraging Optuna's default Tree-structured Parzen Estimator (TPE) algorithm to balance exploration and exploitation effectively.

To mitigate the impact of random sampling during training, 10-fold cross-validation was used to evaluate the model's performance. It is worth noting that Test 1# and Test 2# were not involved in the cross-validation process; they were reserved for the final evaluation of model performance.

5. Model performance presentation

5.1. Damage states assessment

With reference to the provisions of HAZUS [83] and Pagni et al. [84], damage states of LRBs are divided into three categories corresponding to the green, yellow and red labeling categories, respectively: 1. **Non-slipping (NS)**: When the LRB does not slip at all during an earthquake, it is recorded as non-slipping (NS) state (corresponding to no damage in HAZUS); 2. **Bearing-sliding (BS)**: When the LRB slides, but the sliding displacement does not exceed 250 % of the horizontal shear strain, it is recorded as the bearing-sliding state. This is associated with "slight," "moderate," and "extensive" damage categories in HAZUS; 3. **Bearing Failure (BF)**: When the sliding displacement of LRB is greater than 250 %, it is recorded as bearing failure (BF) state (corresponding to complete damage in HAZUS). This categorization provides a more transparent presentation of the model's performance.

Cross-entropy is a commonly used metric to measure the difference between probability distributions. In this study, cross-entropy was used as the loss function to train the classification model. Cross-entropy is defined as follows:

$$\text{Cross} - \text{Entropy} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log(\hat{y}_{ic}) \tag{4}$$

where, N is sample size, C is number of categories, y_{ic} is realistic distribution for sample i, $\hat{y}_{ic}$ is the predicted probability of sample i belonging to category c.

A confusion matrix is a table used to assess the performance of classification algorithms. Fig. 8 presents the confusion matrix of the best model obtained through cross-validation for each model, representing their peak performance. For instance, in Fig. 8 **(a)**, the TabNet correctly predicts 1368 green labels, 1155 yellow labels and 890 red labels. In the confusion matrix, the blue grid in the bottom-right corner represents the accuracy, while the grey grids along the right and bottom represent the recall and precision for each category, respectively. Accuracy refers to the proportion of total entries that were correctly predicted, reflecting the model's overall performance. Additionally, performance can be assessed in more detail using recall and precision: Recall measures the proportion of actual instances of each category that were correctly

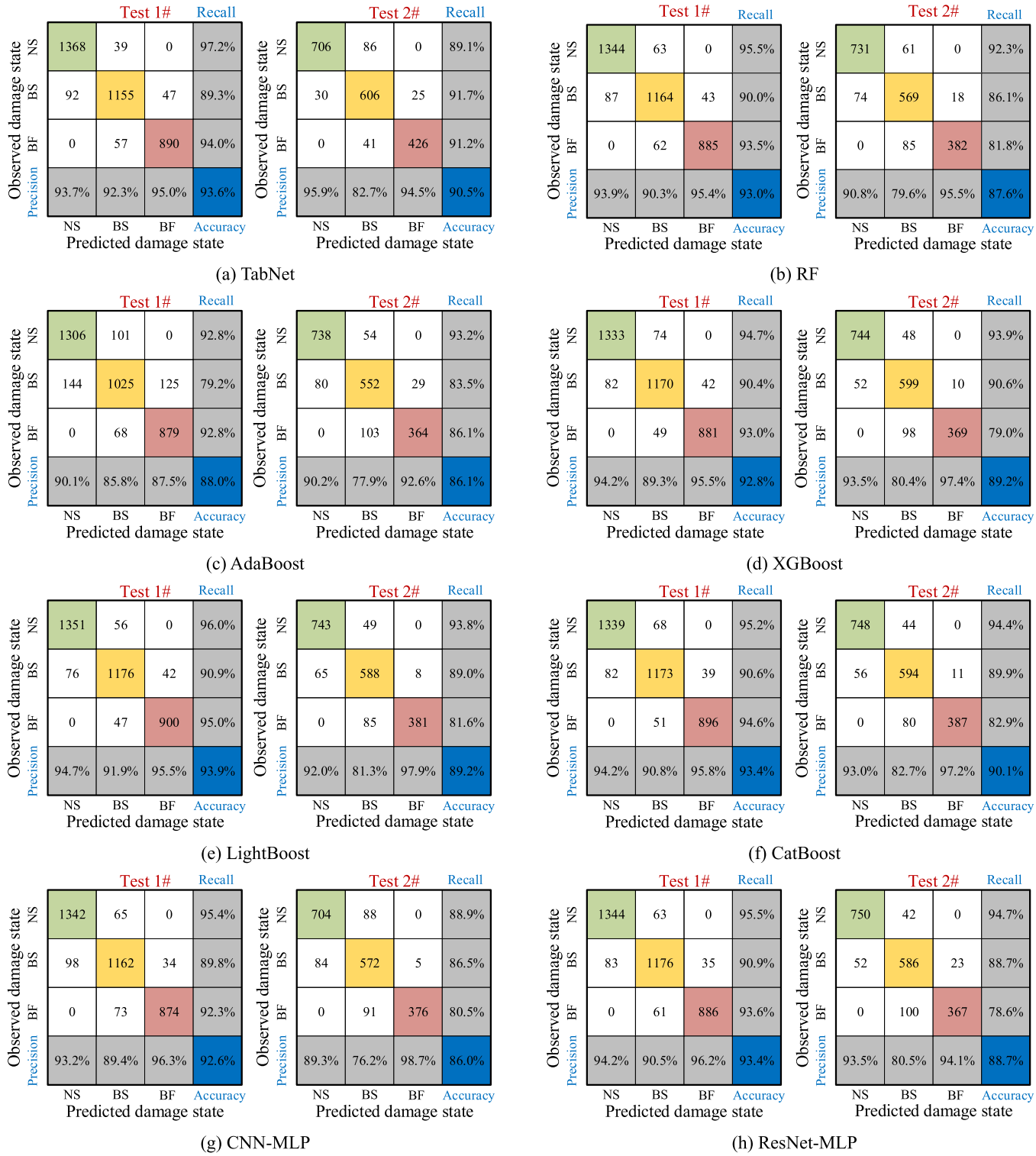


Fig. 8. Confusion matrix of various ML models.

predicted; Precision evaluates the proportion of predicted instances for each category that were correctly classified.

As shown in Fig. 8, the accuracy of all the models on the test sets exceeded 86 %. This indicates that the proposed prediction framework, both the tree-based ensemble models and DNNs, achieved satisfactory prediction results in the ULRBs-supported continuous girder bridges with TYSD. The peak accuracy of all ML models (except AdaBoost) on Test 1# did not differ much and was above 91.0 %. Among them, LightBoost (93.9 %) achieved the highest accuracy on Test 1#, followed by TabNet (93.6 %). AdaBoost, on the other hand, performed the worst on Test 1# with an accuracy of only 88.0 %.

On Test 2#, TabNet took the lead with an excellent accuracy of 90.5 %, while CNN-MLP (86.0 %) and AdaBoost (86.1 %) exhibited the worst performances. TabNet demonstrated its capability to accurately predict damage in structures subjected to unknown earthquakes, showcasing its ability to learn the relationship between features and

target categories.

It was noteworthy that, although LightBoost (89.2 %) and CatBoost (90.1 %) achieved strong accuracy on Test 2#, the recall scores for the red labels (bearing failure, BF) were relatively low, at 81.6 % and 82.9 %, respectively. In contrast, TabNet maintained a high recall value for the red label on Test 2# at 91.2 %, indicating superior accuracy in predicting the unseating of the bridge.

Almost all of ResNet-MLP's performance on Test 1# and Test 2# was superior to those of CNN-MLP, suggesting that the modifications to CNN-MLP were effective. Both ResNet-MLP (92.6 %) and CNN-MLP (93.4 %) performed well on Test 1#, but their peak performance on Test 2# was only slightly better than that of AdaBoost and RF, and generally inferior to most tree-based ensemble methods and TabNet.

Notably, the precision for the yellow labels was slightly lower compared to the other two categories. This phenomenon was likely due to the fact that the yellow label category was not as distinctly separated from both the red and green label categories, increasing the likelihood of misclassification. This resulted in a higher number of false positives for yellow labels.

Additionally, the average performance and stability of each model in 10-fold cross-validation were presented in Table 3. In addition to accuracy, recall, and precision, F1-score and the area under the receiver operating characteristic curve (AUC) were included to provide a more comprehensive evaluation of model performance. Models with the best average performance were highlighted with bold lines, while those with the second-best performance were underlined.

In general, almost all tree-based ensemble models, except AdaBoost, were more stable than DNNs. TabNet achieved the highest average scores in most metrics on Test 1# and Test 2#, with the exception of AUC and precision, where it was slightly outperformed by CatBoost and LightBoost. CatBoost had quite high AUC scores on both test sets, and was second only to TabNet in the rest of the evaluation metrics, while it also demonstrated excellent stability. This indicated that CatBoost performed well in classification tasks with all possible thresholds and had robustness to unbalanced data classes.

Overall, CatBoost and LightBoost offered excellent stability and efficiency, making them strong choices for many tasks. However, if higher evaluation metrics such as accuracy and recall were the primary goals, TabNet stood out as the superior option. TabNet outperformed the other methods in terms of both accuracy on Test 1# and Test 2#, as well as the precision and recall for the red labels. Accurate prediction of the red labels is particularly critical in post-disaster scenarios, as it enables emergency responders to make informed decisions in a timely manner.

5.2. Seismic response prediction

Accurate prediction of the displacement response of sliding bearings is critical for the reinforcement design and risk assessment of ULRB-supported bridges during earthquakes. In this subsection, bridges in the BS (bearing-sliding) state, as described previously, were selected for the dataset. For this task, Huber Loss was chosen as the loss function to train the regression model. Huber Loss combines the advantages of both Mean Squared Error (MSE) and Mean Absolute Error (MAE), reducing the influence of outliers while maintaining good numerical stability. The Huber Loss function is defined as:

$$L_\delta(y-\hat{y}) = \begin{cases} \frac{1}{2}(y-\hat{y})^2, |y-\hat{y}| \le \delta \\ \delta|y-\hat{y}| - \frac{1}{2}\delta^2, |y-\hat{y}| > \delta \end{cases} \tag{5}$$

where y and $\hat{y}$ represents the actual and predicted values of the sample, respectively, and δ is a hyper-parameter to control the switching point. In order to validate the model's generalization performance, it is essential to examine its performance in the test set. The coefficient of determination R^2 is used to evaluate its performance on the two sets:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2} \tag{6}$$

Fig. 9 shows the scatter plots of various ML models on Test 1# and Test 2#, with R^2 values and 25 % error lines annotated. It could be observed that the R^2 values of almost all ML models (except AdaBoost) were above 0.9, indicating that this prediction framework was also applicable for regression task of seismic response. The performance of ResNet-MLP (0.954) and TabNet (0.953) on Test 1# was noteworthy, however, on Test 2#, it is comparable to RF, XGBoost, LightBoost and CNN-MLP, with a performance of 0.91, which was not as good as that of CatBoost (0.925) and ResNet-MLP (0.922). The ResNet-MLP demonstrated excellent performance in both Test 1# and Test 2#.

Similarly, the average performance and stability of each model in cross-validation is shown in Table 4. R^2 visually demonstrates the goodness of fit of the data. It is intuitive to show how well the data fit, but may be unstable in case the data are unbalanced or contain outliers. Therefore, MAE, MSE, mean relative error (MRE) and standard deviation of relative errors (SDRE) were added to describe the dispersion and stability of the models. As shown in Table 4, TabNet and ResNet-MLP maintained excellent average performance on Test 1#, but their performance on Test 2# was not as good as most of the tree-based ensemble models. CatBoost exhibited excellent average performance on both Test 1# and Test 2#, followed by LightBoost, which also demonstrated high

Table 3
Average performance and stability of various ML models in damage states assessment.

ML model		Test 1# (%)					Test 2# (%)				
		Accuracy	Precision	Recall	F1-score	AUC	Accuracy	Precision	Recall	F1-score	AUC
DNN	TabNet	**93.3 ± 0.34**	**93.4 ± 0.35**	**93.4 ± 0.33**	**93.4 ± 0.33**	99.0 ± 0.09	**89.2 ± 0.48**	90.0 ± 0.38	**88.4 ± 0.69**	**89.0 ± 0.57**	97.9 ± 0.18
	CNN-MLP	91.3 ± 0.77	91.5 ± 0.77	91.6 ± 0.75	91.5 ± 0.75	92.9 ± 2.63	84.8 ± 1.25	86.3 ± 1.35	84.4 ± 1.48	85.0 ± 1.38	89.4 ± 2.22
	ResNet-MLP	91.8 ± 1.07	91.9 ± 1.07	91.9 ± 0.99	91.9 ± 1.03	93.2 ± 2.62	86.6 ± 0.91	87.1 ± 0.98	85.8 ± 0.83	86.3 ± 0.82	91.0 ± 2.42
Ensemble Learning	RF	92.7 ± 0.15	92.9 ± 0.17	92.8 ± 0.17	92.8 ± 0.17	99.0 ± 0.02	86.6 ± 0.31	88.0 ± 0.30	85.3 ± 0.38	86.2 ± 0.35	97.2 ± 0.12
	AdaBoost	86.6 ± 1.02	86.6 ± 0.91	86.9 ± 0.97	86.7 ± 0.94	81.2 ± 0.75	85.6 ± 1.10	86.3 ± 1.02	84.8 ± 1.11	85.4 ± 1.04	84.0 ± 2.05
	XGBoost	92.5 ± 0.16	92.7 ± 0.15	92.5 ± 0.16	92.6 ± 0.15	98.9 ± 0.02	88.1 ± 0.38	89.5 ± 0.30	87.9 ± 0.49	87.9 ± 0.43	97.2 ± 0.06
	CatBoost	93.1 ± 0.17	93.3 ± 0.17	93.2 ± 0.16	93.3 ± 0.17	**99.1 ± 0.02**	88.8 ± 0.76	**90.1 ± 0.63**	87.5 ± 0.88	88.4 ± 0.81	**98.0 ± 0.15**
	LightBoost	92.8 ± 0.13	92.9 ± 0.13	92.8 ± 0.13	92.8 ± 0.13	99.0 ± 0.01	88.9 ± 0.27	90.0 ± 0.23	87.9 ± 0.30	88.8 ± 0.27	97.5 ± 0.04

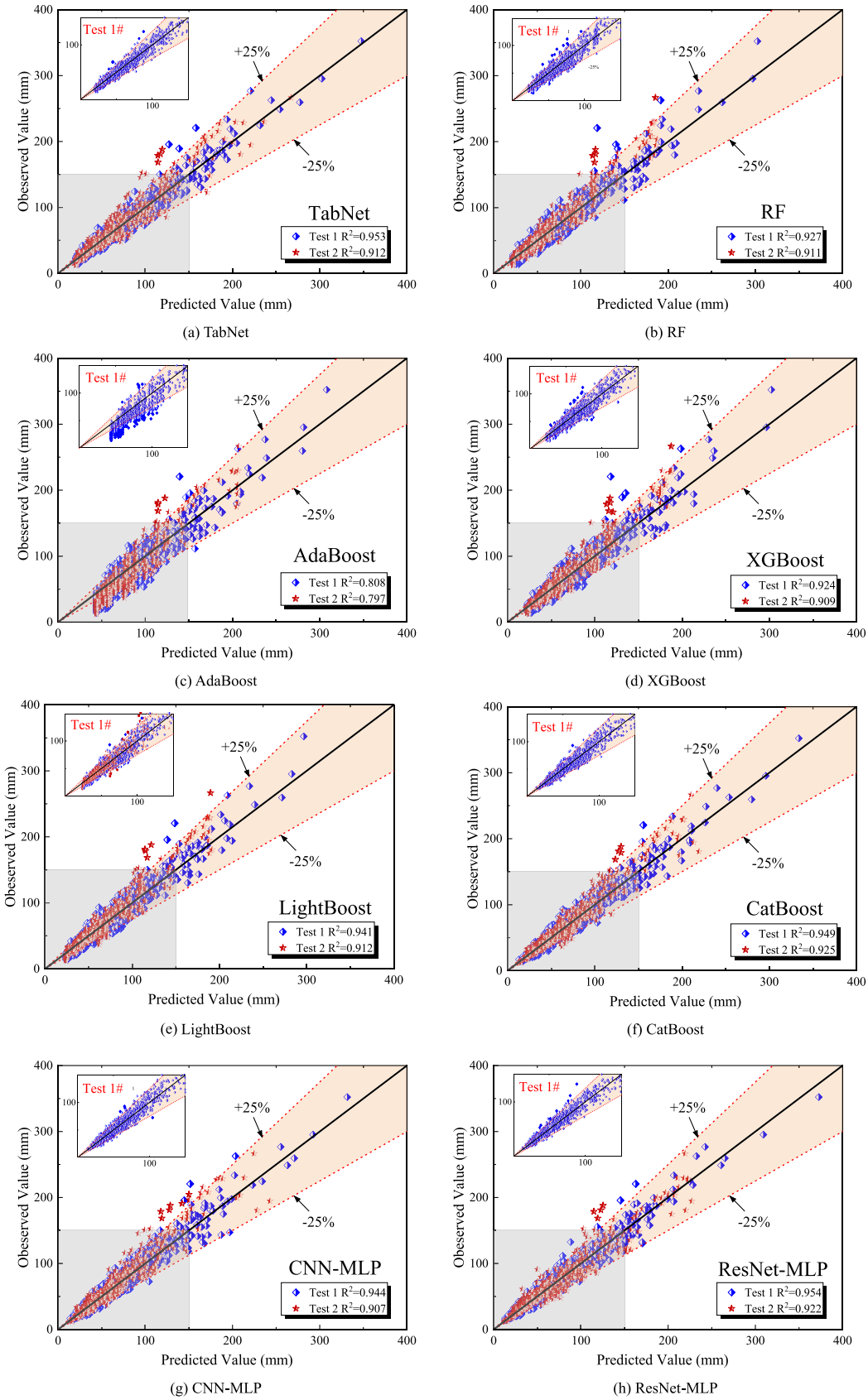


Fig. 9. Distribution of the predicted and true values of various ML models.

Table 4
Average performance and stability of various ML models in seismic response prediction ($\times 10^{-2}$).

ML model		Test 1#					Test 2#				
		R^2	MAE	MSE	MRE	SDRE	R^2	MAE	MSE	MRE	SDRE
DNN	TabNet	**94.8 ± 0.63**	**0.52 ± 0.040**	0.006 ± 0.0007	**−0.002 ± 0.044**	4.52 ± 0.043	89.3 ± 1.15	0.70 ± 0.029	0.013 ± 0.0020	0.20 ± 0.080	4.6 ± 0.054
	CNN-MLP	93.4 ± 0.88	2.04 ± 0.133	0.08 ± 0.0108	0.15 ± 0.205	2.80 ± 0.200	89.4 ± 0.67	2.40 ± 0.076	0.15 ± 0.009	0.43 ± 0.197	3.80 ± 0.132
	ResNet-MLP	94.4 ± 0.83	1.88 ± 0.131	0.07 ± 0.0106	−0.06 ± 0.487	2.62 ± 0.201	90.0 ± 1.56	2.43 ± 0.186	0.14 ± 0.022	−0.38 ± 0.643	3.66 ± 0.242
Ensemble Learning	RF	92.3 ± 0.18	0.62 ± 0.005	0.008 ± 0.0002	−0.02 ± 0.008	0.91 ± 0.011	90.8 ± 0.19	0.65 ± 0.066	0.01 ± 0.0002	0.21 ± 0.025	1.02 ± 0.011
	AdaBoost	79.7 ± 1.02	1.24 ± 0.036	0.02 ± 0.00110	−0.90 ± 0.042	1.16 ± 0.020	79.8 ± 0.57	1.29 ± 0.028	0.024 ± 0.0008	−0.76 ± 0.643	1.35 ± 0.024
	XGBoost	92.1 ± 0.18	0.63 ± 0.004	0.008 ± 0.0002	−0.03 ± 0.006	0.92 ± 0.010	91.1 ± 0.42	0.65 ± 0.014	0.01 ± 0.0005	0.16 ± 0.028	1.01 ± 0.025
	CatBoost	94.7 ± 0.10	0.53 ± 0.004	**0.006 ± 0.0001**	−0.02 ± 0.005	**0.75 ± 0.007**	**92.1 ± 0.40**	**0.63 ± 0.016**	**0.009 ± 0.0005**	**0.06 ± 0.035**	**0.96 ± 0.025**
	LightBoost	93.9 ± 0.10	0.57 ± 0.003	0.006 ± 0.0001	−0.04 ± 0.006	0.80 ± 0.006	91.2 ± 0.35	0.65 ± 0.007	0.01 ± 0.0004	0.12 ± 0.019	1.01 ± 0.019

stability. The peak performance of ResNet-MLP was quite good, however, the variance of its performance metrics was significantly higher compared to tree-based ensemble models and TabNet.

In summary, ResNet-MLP exhibited the best peak performance on both Test 1# and Test 2#. This makes ResNet-MLP an appropriate choice for application scenarios in anomaly detection or early warning systems, where a single serious error can lead to unacceptable consequences. Therefore, it is crucial to ensure the model's reliability and accuracy under extreme conditions. In terms of stability and efficiency, CatBoost achieved the highest overall score and also demonstrated impressive peak performance. CatBoost is undoubtedly a better choice for common scenarios that do not require relatively high accuracy in prediction. Although TabNet performed well on Test 1# and was suitable for tasks such as post-earthquake damage assessment, its average performance, extreme performance, and stability in the face of unknown ground shaking were not as impressive as CatBoost.

5.3. Computational complexity and training efficiency

The computational complexity and training efficiency of models serve as significant reference indicators for model selection in practical engineering applications. The computational complexity metrics for DNNs include the number of parameters and Multiply–Accumulate Operations (MACs), while ensemble models are assessed by the number of trees. The training efficiency was assessed by the training time per under consistent environments. The computational complexity and training efficiency metrics for each model are shown in Table 5.

The results indicated that tree-based ensemble methods exhibited low computational complexity and short training times. This characteristic made them suitable for rapid training and efficient deployment. Conversely, DNN models were more complex and require more computational resources and time costs. TabNet achieved a balance between model complexity and training cost, while preserving excellent and robust performance. Consequently, it is particularly suitable for application scenarios that necessitate both interpretability and accuracy.

Table 5
Computational complexity and training efficiency of various ML models.

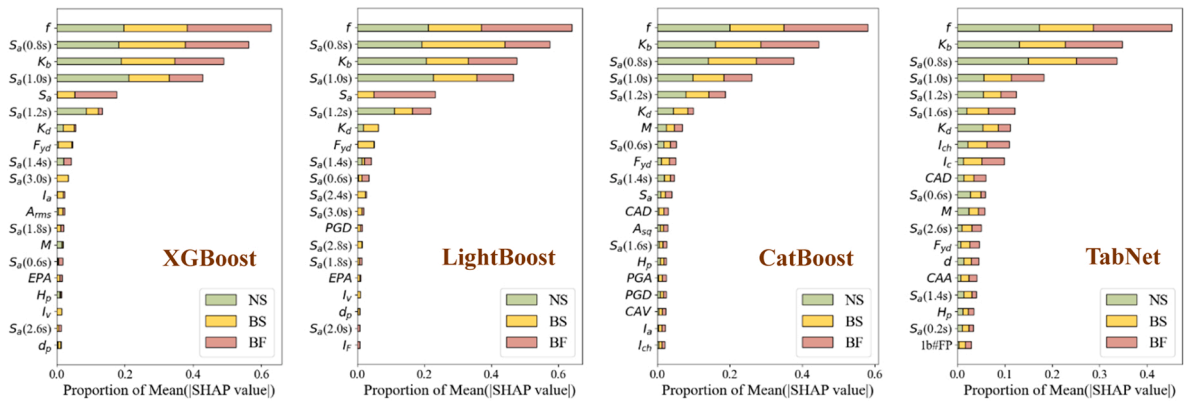
Models	Parameters	MACs	Number of trees	Training time per Epoch (s)	Total training time (min)
TabNet	22.06k	523.12k	—	1.03	0.98
CNN-MLP	243.62k	19.15 M	—	0.92	2.58
ResNet-MLP	1.07 M	129.19 M	—	2.20	5.42
RF	—	—	445	—	0.17
AdaBoost	—	—	199	—	1.01
XGBoost	—	—	305	—	0.07
CatBoost	—	—	868	—	0.79
LightBoost	—	—	303	—	0.027

5.4. Effect of input variables on predicting the damage states based on ML models

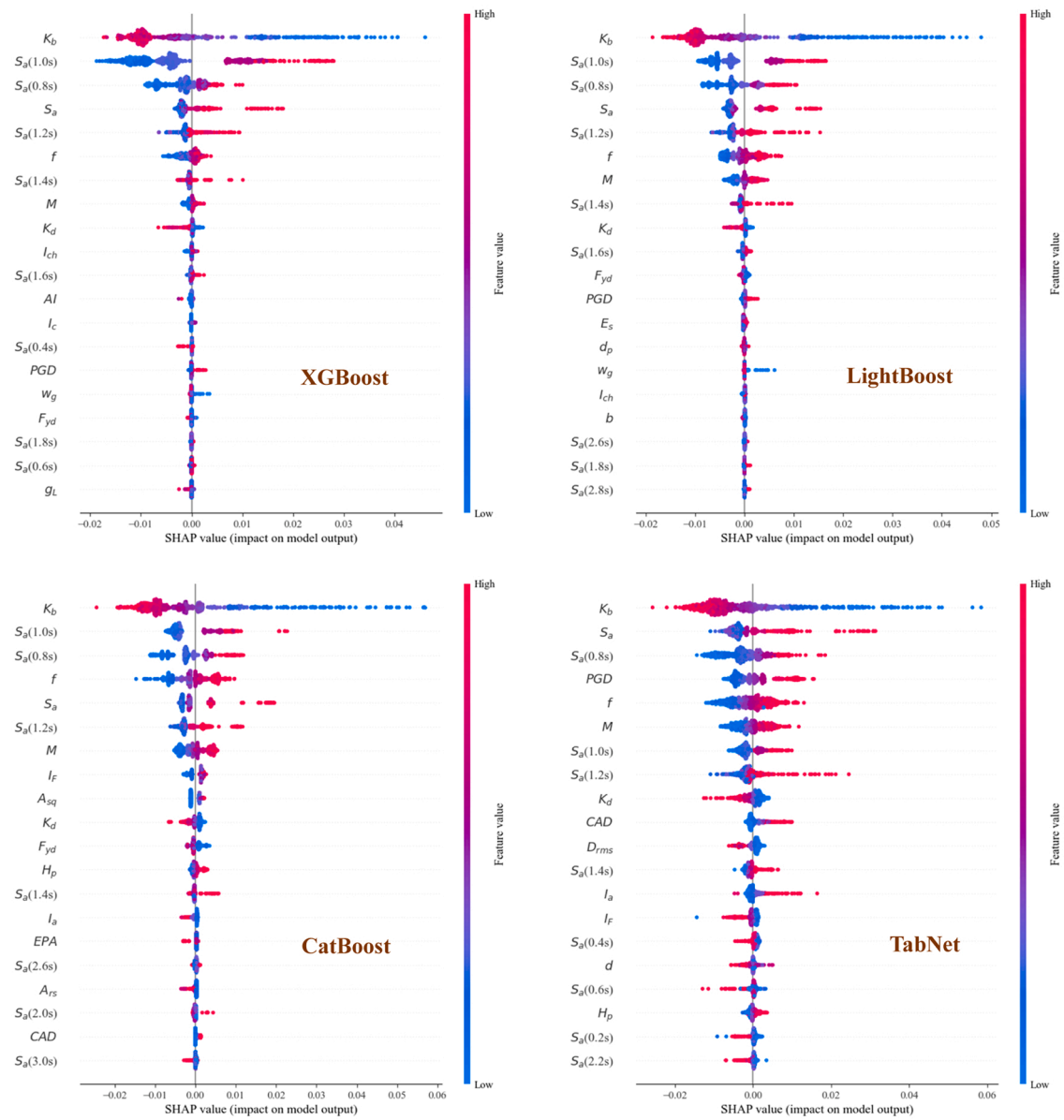
To mitigate bias in the ML models, this subsection utilizes the SHAP (SHapley Additive exPlanations) algorithm [85] to extract importance scores for each input variable, thereby enhancing the interpretability of the models. SHAP is grounded in Shapley value theory from cooperative game theory, which quantifies the marginal contribution of each feature in various feature combinations. By averaging these marginal contributions, SHAP provides a clear explanation of how individual features influence the predicted outputs. Compared to methods such as LIME [86] and Integrated Gradients [87], SHAP has a stronger theoretical foundation [85] and delivers more stable explanatory results.

This subsection focuses on the models (TabNet, CatBoost, LightBoost, and XGBoost) that performed well in previous sections, excluding ResNet-MLP from the discussion. Although ResNet-MLP demonstrated excellent peak performance in seismic response prediction, CatBoost achieved comparable peak performance while offering better stability and more consistent average performance. Moreover, the input to deep neural networks (DNNs) like ResNet-MLP consists of time-series data, and the importance scores in this context correspond to values at each time point. As a result, the feature importance of time-series data cannot be directly compared to the feature importance derived from tabular data used in the other models.

Fig. 10 **(a)** shows the top 20 feature importance scores for each ML models in the damage states assessment. Fig. 10 **(b)** shows the importance and trend of input variables for each ML model in structural response prediction. The vertical axis represents the order of important input variables from top to bottom, and the horizontal axis indicates the probability of the predicted results. Positive (negative) values indicate an increase (decrease) in the probability of prediction results. It can be observed that friction coefficient (f), bearings elastic stiffness (K_b), and acceleration spectra ($S_a(0.8s)$ & $S_a(1.0s)$) consistently demonstrated significant influence in all ML models. The friction coefficient (f) directly affects the sliding performance of the laminated-rubber bearings. A lower coefficient can provide better seismic isolation but will increase the displacement of the girder, while a higher coefficient can limit the displacement but will increase the response demands of the substructure. Similarly, a higher elastic stiffness of the bearing (K_b) can effectively limit excessive deformation, enhancing the stability of the superstructure, yet excessively high stiffness may reduce the seismic



(a)



(b)

Fig. 10. Relative importance of input variables on (a)damage states; (b) seismic response prediction.

isolation effectiveness. $S_a(0.8s)$ and $S_a(1.0s)$ are closely related to the dynamic characteristics of small-and-medium-span continuous girder bridges. These spectral accelerations correspond to the seismic demand that is representative of fundamental vibration modes of typical bridges to predict seismic responses and potential damage. The maximum sliding displacements of the bearings exhibited a decreasing trend with an increase in K_b, and an increasing trend with an increase in $S_a(0.8s)$ and $S_a(1.0s)$. This is in accordance with the principles of mechanics, indicating that the model's decisions are reasonable and reliable.

It is worth noting that there are some differences between the decision-making processes of TabNet and Boosting methods. For example, in Fig. 10 **(a)**, TabNet placed greater emphasis on features such as I_{ch} and I_c of the GM records, which were not prominent in the Boosting methods. In XGBoost and LightBoost, it demonstrated that S_a exerted a significant influence on the BS and BF classes, particularly the BF class. In contrast, in TabNet, S_a was of relatively minor importance. The reason for the above phenomenon is that tree-based Boosting methods improve the model performance by building a series of weak learners (DT) and progressively refining them [88]. In contrast, TabNet uses an attention mechanism to select important features at each decision step and makes decisions through layer-wise feature selection and processing [78]. This dynamic feature selection is based on weight adjustments within each network layer, offering a more adaptive and nonlinear evaluation of feature importance. However, it also results in suboptimal performance on some simple datasets [78,89], such as in the seismic response prediction of sliding bearings in Section 5.2.

To further elucidate how TabNet's decision-making approach fundamentally differs from both boosting methods and conventional DNNs, the t-distributed Stochastic Neighbor Embedding (t-SNE) method was utilized to visualize the high-dimensional feature embeddings learned by TabNet, as shown in Fig. 11. Additionally, heatmaps illustrating average feature masks across the five decision steps in TabNet (Fig. 12) were generated, highlighting the dynamically identified importance of features at each decision step. These visualizations effectively illustrated TabNet's distinctive capability of adaptive feature selection, clearly distinguishing it from other considered models.

6. Discussion

The methods in this study are effective for rapid damage states assessment and seismic response prediction of bridge, but some limitations should be recognized. The primary limitation is attributable to reliance on FE simulations, which may not capture all aspects of real-world structural behavior and earthquake events. In subsequent research, the effectiveness of the model can be further enhanced by incorporating actual structural monitoring data from seismic events.

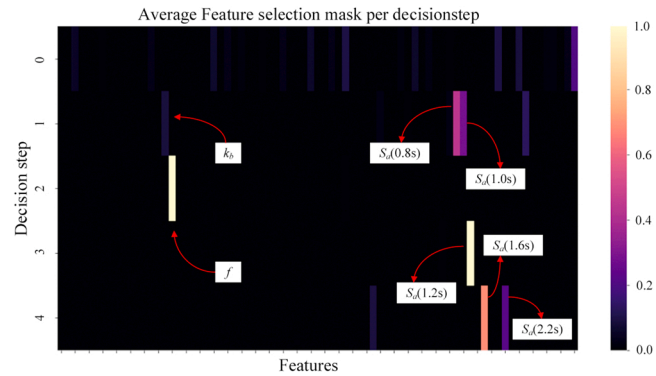
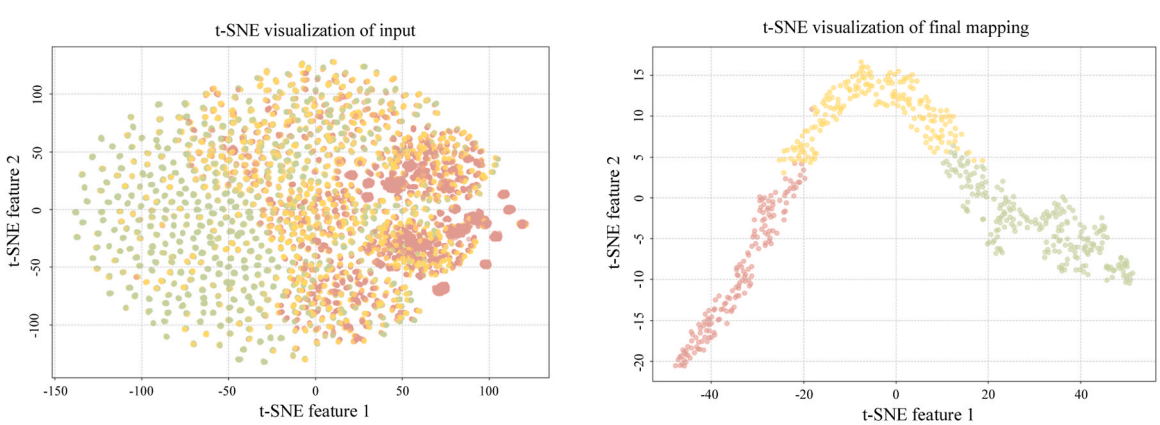


Fig. 12. The aggregate feature importance mask.

Additionally, despite the collection of extensive information regarding bridge structures and earthquake events, the dataset remains incapable of comprehensively representing all potential scenarios. The effectiveness of ML models heavily depends on the data quality. The development of a more comprehensive dataset will substantially enhance the model's robustness and generalization.

A further challenge is the practical applicability of ML models. A complex model can lead to increased costs during training, as well as subsequent challenges during debugging and maintenance. This undoubtedly makes deployment in engineering more challenging. The simplification of the model architecture or the development of a user-friendly interface has the potential to significantly facilitate the practical application of these predictive models.

In terms of feature selection, the data types that are conducive to ensemble models and DNN models are distinct. Ensemble models have been shown to be effective in the handling of manually extracted tabular features. However, these models are heavily dependent on prior domain knowledge, which may result in the neglect of critical features present in the time or frequency domains of GM records. In contrast, DNN models have the advantage of automatically extracting features directly from GM records. Nonetheless, this capability significantly increases model complexity and computational cost. Furthermore, the performance of DNN models is constrained by the quantity and quality of the data, potentially resulting in limited effectiveness when dealing with small-scale or low-quality datasets. Future studies could benefit from hybrid approaches, combining domain-specific feature engineering with automated feature extraction to improve overall predictive performance.



(a) t-SNE visualization of input (b) t-SNE visualization of final mapping

Fig. 11. t-SNE visual analysis of various levels of features.

7. Conclusion

This study presents a rapid damage states assessment and seismic response prediction framework for ULRB-supported bridges based on ML algorithms and FE simulations. The study explores the impact of various input data types and ML models on prediction performance. Specifically, three DNNs and five ensemble models were considered, including TabNet, a novel DNN architecture optimized for tabular data. A total of 38,400 samples were generated by FE simulation for ML models development. These FE models account for geometric and material uncertainties, representing a diverse range of practical bridge designs. The performance of the eight ML models was evaluated through damage states classification and seismic response regression tasks, with consideration given to both peak and average performance, as well as model stability. Furthermore, SHAP was used to visualize the feature importance scores for each ML model. The main conclusions drawn from this study are as follows:

1. All selected ML models, except AdaBoost, demonstrated effective capabilities in predicting both damage states and seismic responses, confirming the success of the proposed prediction framework. Tree-based ensemble models exhibited higher average performance and greater stability, while DNN architectures showed superior peak performance. Notably, no single model outperformed others across all metrics and tasks.
2. TabNet is recommended for damage state assessment based on its superior prediction performance. TabNet achieved the highest accuracy across most prediction metrics in both test sets, particularly for predicting the most severe damage state (bearing failure, BF). On Test 1# and Test 2#, TabNet achieved peak accuracies of 93.6 % and 90.5 %, respectively, with average accuracies of 93.3 % and 89.2 %. In comparison, CatBoost ranked second with peak accuracies of 93.4 % and 90.1 %, and average accuracies of 93.1 % and 88.8 %.
3. For seismic responses prediction, CatBoost is recommended due to its excellent balance of performance, stability, and efficiency. Although ResNet-MLP exhibited similar peak performance to CatBoost (0.954 on Test 1# and 0.922 on Test 2#), it is more computationally expensive and less stable, with average performances of 0.944 and 0.900, respectively. In contrast, CatBoost achieved an average performance of 0.947 and 0.921. While TabNet showed satisfactory results on Test 1#, its peak performance (0.912) and average performance (0.895) on Test 2# were not as competitive as CatBoost.
4. SHAP effectively identified important parameters affecting damage states assessment and seismic response predictions. The results revealed that key parameters such as the friction coefficient (f), bearings elastic stiffness (K_b), and acceleration spectra ($S_a(0.8s)$ & $S_a(1.0s)$) were consistently influential across all models. In particular, K_b, $S_a(0.8s)$, and $S_a(1.0s)$ demonstrated significant impacts on the seismic response, aligning with mechanical principles, which confirms the model's reliability.

In summary, TabNet is highly effective for damage state classification, while CatBoost provides the best balance of stability, efficiency, and performance for seismic response prediction. Both models demonstrated strong potential for real-world applications in bridge performance assessment and seismic risk evaluation.

In addition, the framework proposed in this study cannot accurately predict residual displacements. This may be due to an insufficient number of datasets, a lack of diversity in GMs, or the inability of ML models to capture the spectral characteristics and global information of GM records. The authors will conduct research on predicting the residual displacements of related structures in the future.

CRediT authorship contribution statement

Bian Jiachen: Writing – review & editing, Visualization, Validation. **Guo Tong:** Writing – review & editing, Visualization, Validation. **Gu Ranxing:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Jia Junfeng:** Writing – review & editing, Validation, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of Competing Interest

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

Acknowledgments

This work was supported by the International Science & Technology Cooperation Program of China (Grant No. 2024YFE0198400), National Natural Science Foundation of China (Grant No. 52178449). The opinions, findings, conclusions, and recommendations presented in this study are those of the authors and do not necessarily reflect the view of the sponsors.

Appendix

Appendix Table A
Detailed hyperparameters, search spaces, distributions, and descriptions for each model optimized using Optuna

Models	Hyperparameters	Search Space	Distribution	Description
TabNet	n_d	[8,16]	IntUniform	Width of the decision prediction layer.
	n_a	[8,16]	IntUniform	Width of the attention embedding for each mask.
	n_steps	[3,10]	IntUniform	Number of steps in the architecture.
	gamma	[1.0, 2.0]	FloatUniform	This is the coefficient for feature reusage in the masks.
	n_independent	[1,5]	IntUniform	Number of independent Gated Linear Units layers at each step.
	n_shared	[1,5]	IntUniform	Number of shared Gated Linear Units at each step.
	momentum	[0.01, 0.4]	FloatUniform	Momentum for batch normalization.
	lambda_sparse	[1e–5]	LogUniform	Extra sparsity loss coefficient.
	learning_rate	[1e–5]	LogUniform	Initial learning rate.
	step_size	[10,20]	IntUniform	Number of epochs between each learning rate update.
	gamma	[0.7, 0.99]	FloatUniform	Multiplicative factor applied to the learning rate at each step.
CNN-MLP	in_channel	{8, 16, 32, 64}	Categorical	Number of channels in the initial convolutional layer.
	in_kernel_size	{3, 5, 7}	Categorical	Kernel size of the initial convolutional layer.

(*continued on next page*)

Appendix Table A (*continued*)

Models	Hyperparameters	Search Space	Distribution	Description
	GM_Info	{64, 128, 256}	Categorical	Dimension of the GM information input to the decoder.
	learning_rate	[1e–5]	LogUniform	Initial learning rate.
	step_size	[10,20]	IntUniform	Number of epochs between each learning rate update.
	gamma	[0.7, 0.99]	FloatUniform	Multiplicative factor applied to the learning rate at each step.
	in_channel	{8, 16, 32, 64}	Categorical	Number of channels in the initial convolutional layer.
	in_kernel_size	{3, 5, 7}	Categorical	Kernel size of the initial convolutional layer.
ResNet-MLP	GM_Info	{64, 128, 256}	Categorical	Dimension of the GM information input to the decoder.
	learning_rate	[1e–5]	LogUniform	Initial learning rate.
	step_size	[10,20]	IntUniform	Number of epochs between each learning rate update.
	gamma	[0.7, 0.99]	FloatUniform	Multiplicative factor applied to the learning rate at each step.
RF	n_estimators	[50,500]	IntUniform	Number of decision trees in the forest.
	max_depth	[3,50]	IntUniform	The maximum depth allowed for each tree.
	min_samples_split	[2,8]	IntUniform	Minimum number of samples required to split an internal node.
	min_samples_leaf	[1,5]	IntUniform	The minimum number of samples required at a leaf node.
AdaBoost	n_estimators	[20,200]	IntUniform	Number of weak learners to be used in the boosting ensemble
	max_depth	[1,10]	IntUniform	The maximum depth allowed for each tree
	learning_rate	[1e−3, 1.0]	LogUniform	The weight applied to the contribution of each learner.
XGBoost	n_estimators	[50,500]	IntUniform	Number of boosting rounds
	max_depth	[3,10]	IntUniform	The maximum depth allowed for each tree.
	learning_rate	[1e−3, 1.0]	FloatUniform	The step size shrinkage used to prevent overfitting
CatBoost	iterations	[500,1000]	IntUniform	The total number of boosting iterations
	depth	[3,16]	IntUniform	The maximum depth allowed for each tree.
	learning_rate	[1e−3, 0.1]	LogUniform	Step size shrinkage used in updating the weights.
LightBoost	n_estimators	[50,500]	IntUniform	Number of trees to be used in the boosting ensemble
	max_depth	[3,20]	IntUniform	The maximum depth of each tree
	learning_rate	[1e−3, 0.1]	LogUniform	The shrinkage rate applied at each boosting iteration

References

[1] Lin X, Zhang H, Cheng H, et al. Field investigation on severely damaged aseismic buildings in 2014 Ludian earthquake. Earthq Eng Eng Vib 2015;14:169–76. https://doi.org/10.1007/s11803-015-0014-5.

[2] Wald D, Lin K-W, Porter K, et al. ShakeCast: automating and improving the use of shakemap for post-earthquake decision-making and response. Earthq Spectra 2008; 24(2):533–53. https://doi.org/10.1193/1.2923924.

[3] Mangalathu S. Performance based grouping and fragility analysis of box-girder bridges in California. Doctoral dissertation. Georgia Institute of Technology; 2017. ⟨http://hdl.handle.net/1853/58291⟩.

[4] Mangalathu S, Heo G, Jeon J-S. Artificial neural network based multi-dimensional fragility development of skewed concrete bridge classes. Eng Struct 2018;162: 166–76. https://doi.org/10.1016/j.engstruct.2018.01.053.

[5] Mangalathu S, Jeon J. Ground motion-dependent rapid damage assessment of structures based on wavelet transform and image analysis techniques. J Struct Eng 2020;146(11):04020230. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002793.

[6] ASCE/SEI 7-22. Minimum design loads and associated criteria for buildings and other structures. Reston, Virginia: ASCE; 2022.

[7] Mangalathu S, Jeon J-S, Jiang J. Skew adjustment factors for fragilities of California box-girder bridges subjected to near-fault and far-field ground motions. J Bridge Eng 2019;24(1):04018109. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001338.

[8] Cioffi R, Travaglioni M, Piscitelli G, et al. Artificial intelligence and machine learning applications in smart production: progress, trends, and directions. Sustainability 2020;12(2):492. https://doi.org/10.3390/su12020492.

[9] Afshar A, Nouri G, Ghazvineh S, et al. Machine-learning applications in structural response prediction: a review. Pract Period Struct Des Constr 2024;29(3): 03124002. https://doi.org/10.1061/PPSCFX.SCENG-1292.

[10] Moein MM, Saradar A, Rahmati K, et al. Predictive models for concrete properties using machine learning and deep learning approaches: a review. J Build Eng 2023; 63:105444. https://doi.org/10.1016/j.jobe.2022.105444.

[11] Fan G, Li J, Hao H, et al. Data driven structural dynamic response reconstruction using segment based generative adversarial networks. Eng Struct 2021;234: 111970. https://doi.org/10.1016/j.engstruct.2021.111970.

[12] Fayaz J, Galasso C. A deep neural network framework for real-time on-site estimation of acceleration response spectra of seismic ground motions. Comput-Aided Civ Infrastruct Eng 2023;38(1):87–103. https://doi.org/10.1111/mice.12830.

[13] Nie T, Wang S, Wang Y, et al. An effective recognition of moving target seismic anomaly for security region based on deep bidirectional LSTM combined CNN. Multimed Tools Appl 2023. https://doi.org/10.1007/s11042-023-14382-5.

[14] Derakhshani A, Foruzan AH. Predicting the principal strong ground motion parameters: a deep learning approach. Appl Soft Comput 2019;80:192–201. https://doi.org/10.1016/j.asoc.2019.03.029.

[15] Yun H, Feng K, Rayhana R, et al. A multidimensional data fusion neural network for damage localization using ultrasonic guided wave. IEEE Trans Instrum Meas 2023;72:2522912. https://doi.org/10.1109/TIM.2023.3300463.

[16] Rautela M, Senthilnath J, Moll J, et al. Combined two-level damage identification strategy using ultrasonic guided waves and physical knowledge assisted machine learning. Ultrasonics 2021;115:106451. https://doi.org/10.1016/j.ultras.2021.106451.

[17] Zhang H, Cheng X, Li Y, et al. Rapid seismic damage state assessment of RC frames using machine learning methods. J Build Eng 2023;65:105797. https://doi.org/10.1016/j.jobe.2022.105797.

[18] Chen S, Feng D, Wang W, et al. Probabilistic machine-learning methods for performance prediction of structure and infrastructures through natural gradient boosting. J Struct Eng 2022;148(8):04022096. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003401.

[19] Xu Y, Lu X, Tian Y, et al. Real-time seismic damage prediction and comparison of various ground motion intensity measures based on machine learning. J Earthq Eng 2020;26(8):4259–79. https://doi.org/10.1080/13632469.2020.1826371.

[20] Rezaei H, Zarfam P, Golafshani EM, et al. Seismic fragility analysis of RC box-girder bridges based on symbolic regression method. Structures 2022;38:306–22. https://doi.org/10.1016/j.istruc.2021.12.058.

[21] Rezaei H, Zarfam P, Golafshani EM, et al. Development of seismic demand prediction models for bridges based on probability approach using symbolic regression method. Comput Struct 2023;282:106991. https://doi.org/10.1016/j.compstruc.2023.106991.

[22] Zhang R, Chen Z, Chen S, et al. Deep long short-term memory networks for nonlinear structural seismic response prediction. Comput Struct 2019;220:55–68. https://doi.org/10.1016/j.compstruc.2019.05.006.

[23] Xia Y, Liu C, Li Y, et al. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. Expert Syst Appl 2017;78:225–41. https://doi.org/10.1016/j.eswa.2017.02.017.

[24] Chen T, Guestrin C. XGBoost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: Association for Computing Machinery; 2016. p. 785–94. https://doi.org/10.1145/2939672.2939785.

[25] Lundberg SM, Erion GG, Lee SI. Consistent individualized feature attribution for tree ensembles. arXiv 2018. https://doi.org/10.48550/arXiv.1802.03888. 1802.03888.

[26] Chen S, Feng D, Wang W, et al. Probabilistic machine-learning methods for performance prediction of structure and infrastructures through natural gradient boosting. J Struct Eng 2022;148(8):04022096. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003401.

[27] Nguyen HD, LaFave JM, Lee YJ, et al. Rapid seismic damage-state assessment of steel moment frames using machine learning. Eng Struct 2022;252:113737. https://doi.org/10.1016/j.engstruct.2021.113737.

[28] Xu JG, Feng DC, Mangalathu S, et al. Data-driven rapid damage evaluation for life-cycle seismic assessment of regional reinforced concrete bridges. Earthq Eng Struct Dyn 2022;51(11):2730–51. https://doi.org/10.1002/eqe.3699.

[29] Zhang H, Cheng X, Li Y, et al. Rapid seismic damage state assessment of RC frames using machine learning methods. J Build Eng 2023;65:105797. https://doi.org/10.1016/j.jobe.2022.105797.

[30] Xu Y, Wei S, Bao Y, et al. Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural

network. Struct Control Health Monit 2019;26(3):e2313. https://doi.org/10.1002/stc.2313.
[31] Yang X, Li H, Yu Y, et al. Automatic pixel-level crack detection and measurement using fully convolutional network. Comput-Aided Civ Infrastruct Eng 2018;33(12):1090–109. https://doi.org/10.1111/mice.12412.
[32] Wu RT, Jahanshahi MR. Deep convolutional neural network for structural dynamic response estimation and system identification. J Eng Mech 2019;145(1):04018125. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001556.
[33] Oh BK, Park Y, Park HS. Seismic response prediction method for building structures using convolutional neural network. Struct Control Health Monit 2020;27(5):e2519. https://doi.org/10.1002/stc.2519.
[34] Mangalathu S, Jeon J. Ground motion-dependent rapid damage assessment of structures based on wavelet transform and image analysis techniques. J Struct Eng 2020;146(11):04020230. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002793.
[35] Wen W, Zhang C, Zhai C. Rapid seismic response prediction of RC frames based on deep learning and limited building information. Eng Struct 2022;267:114638. https://doi.org/10.1016/j.engstruct.2022.114638.
[36] Zhang R, Liu Y, Sun H. Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling. Eng Struct 2020;215:110704. https://doi.org/10.1016/j.engstruct.2020.110704.
[37] Donahue J, Anne, Hendricks L, Guadarrama S, Long-term recurrent convolutional networks for visual recognition and description. In Proceedings of the IEEE conference on computer vision and pattern recognition;2015p. 2625-2634⟨https://openaccess.thecvf.com/content_cvpr_2015/papers/Donahue_Long-Term_Recurrent_Convolutional_2015_CVPR_paper.pdf⟩.
[38] Nie T, Wang S, Wang Y, et al. An effective recognition of moving target seismic anomaly for security region based on deep bidirectional LSTM combined CNN. Multimed Tools Appl 2023. https://doi.org/10.1007/s11042-023-14382-5.
[39] Zhang H, Cheng X, Li Y, et al. Rapid seismic damage state assessment of RC frames using machine learning methods. J Build Eng 2023;65:105797. https://doi.org/10.1016/j.jobe.2022.105797.
[40] Xiang N, Li J. Experimental and numerical study on seismic sliding mechanism of laminated-rubber bearings. Eng Struct 2017;141:159–74. https://doi.org/10.1016/j.engstruct.2017.03.032.
[41] Xiang N, Alam MS, Li J. Yielding steel dampers as restraining devices to control seismic sliding of laminated rubber bearings for highway bridges: analytical and experimental study. J Bridge Eng 2019;24(11):04019103. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001487.
[42] Kwon OS, Jeong SH. Seismic displacement demands on skewed bridge decks supported on elastomeric bearings. J Earthq Eng 2013;17(7):998–1022. https://doi.org/10.1080/13632469.2013.791894.
[43] Han Q, Du X, Liu J, et al. Seismic damage of highway bridges during the 2008 Wenchuan earthquake. Earthq Eng Eng Vib 2009;8(2):263–73. https://doi.org/10.1007/s11803-009-8162-0.
[44] Li J, Peng T, Xu Y. Damage investigation of girder bridges under the Wenchuan earthquake and corresponding seismic design recommendations. Earthq Eng Eng Vib 2008;7(4):337–44. https://doi.org/10.1007/s11803-008-1005-6.
[45] Zhou L, Alam MS, Song A, et al. Probability-based residual displacement estimation of unbonded laminated rubber bearing supported highway bridges retrofitted with Transverse Steel Damper. Eng Struct 2022;272:115053. https://doi.org/10.1016/j.engstruct.2022.115053.
[46] Xiang N, Li J. Effect of exterior concrete shear keys on the seismic performance of laminated rubber bearing-supported highway bridges in China. Soil Dyn Earthq Eng 2018;112:185–97. https://doi.org/10.1016/j.soildyn.2018.04.033.
[47] Shen X, Wang X, Ye Q, et al. Seismic performance of transverse steel damper seismic system for long span bridges. 2017 Eng Struct 2017;141:14–28. https://doi.org/10.1016/j.engstruct.2017.03.014.
[48] Liu C, Gao R, Guo B. Seismic design method analyses of an innovative steel damping bearing for railway bridges. Eng Struct 2018;167:518–32. https://doi.org/10.1016/j.engstruct.2018.04.076.
[49] Wang K., Lu G., Guo W. A study of seismic designs and evaluation methods for small and medium spanning highway girder bridges based on machine learning and earthquake damage experience. 17th World Conference on Earthquake Engineering, 17, WCEE 2020. ⟨https://wcee.nicee.org/wcee/article/17WCEE/2d-0004.pdf⟩.
[50] Deng Y, Ge S, Ge X. Mitigating seismic irregularity of continuous girder bridge with unequal height piers through differential design of shear key strength. Arab J Sci Eng 2023;48:13749–69. https://doi.org/10.1007/s13369-023-07965-9.
[51] Xiang N, Li J. Seismic performance of highway bridges with different transverse unseating-prevention devices. J Bridge Eng 2016;21(9):04016045. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000909.
[52] Zhong HQ, Dang XZ, Deng XW, et al. Seismic performance assessment of girder bridges with composite rubber bearings. J Phys: Conf Ser 2023;2476:012041. https://doi.org/10.1088/1742-6596/2476/1/012041.
[53] Choi E. Seismic analysis and retrofit of mid-America bridges. Georgia Institute of Technology; 2002. ⟨https://www.researchgate.net/profile/Eunsoo-Choi-3/publication/282856959_thesis_all/links/561f745008ae50795b0093ff/thesis-all.pdf⟩.
[54] Shamsabadi A, Yan L. Closed-form force-displacement backbone curves for bridge abutment-backfill systems. Geotech Earthq Eng Soil Dyn IV 2008;2008:1–10. https://doi.org/10.1061/40975(318)159.
[55] Muthukumar S, DesRoches R. A Hertz contact model with non-linear damping for pounding simulation. 2006 Earthq Eng Struct Dyn 2006;35(7):811–28. https://doi.org/10.1002/eqe.557.
[56] Ramanathan KN. Next generation seismic fragility curves for California bridges incorporating the evolution in seismic design philosophy. Doctoral dissertation. Georgia Institute of Technology; 2012. ⟨https://www.proquest.com/docview/1115314245⟩.
[57] Rezaei H, Arabestani S, Akbari R, et al. The effects of earthquake incidence angle on the seismic fragility of reinforced concrete box girder bridges of unequal pier heights. Struct Infrastruct Eng 2020;18(2):1–16. https://doi.org/10.1080/15732479.2020.1842467.
[58] Wang X, Ye A, Ji B. Fragility-based sensitivity analysis on the seismic performance of pile-group-supported bridges in liquefiable ground undergoing scour potentials. Eng Struct 2019;198:109427. https://doi.org/10.1016/j.engstruct.2019.109427.
[59] Brandenberg SJ, Kashighandi P, Zhang J, et al. Fragility functions for bridge in liquefaction-induced lateral spreads. Earthq Spectra 2011;27(3):683–717. https://doi.org/10.1193/1.3610248.
[60] Pang Y, Wu X, Shen G, et al. Seismic fragility analysis of cable-stayed bridges considering different sources of uncertainties. J Bridge Eng 2014;19(4):04013015. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000565.
[61] Barbato M, Gu Q, Conte JP. Probabilistic push-over analysis of structural and soil-structure systems. J Struct Eng 2010;136:1330–41. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000231.
[62] Baker JW, Lin T., Shahi SK, et al. (2011) New ground motion selection procedures and selected motions for the PEER transportation research program. PEER Report, 2011(3). ⟨https://peer.berkeley.edu/sites/default/files/baker_et_al_2011_peer_gm_report.pdf⟩.
[63] Liu PC, Archuleta RJ, Hartzell SH. Prediction of broadband ground-motion time histories: hybrid low/ high frequency method with correlated random source parameters. Bull Seismol Soc Am 2006;96(6):2118–30. https://doi.org/10.1785/0120060036.
[64] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 2012;25(NIPS 2012). ⟨https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf⟩.
[65] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44. https://doi.org/10.1038/nature14539.
[66] Glorot X, Bengio Y, Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics (pp. 249-256). JMLR Workshop and Conference Proceedings;2010. ⟨https://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf⟩.
[67] Khan A, Sohail A, Zahoora U, et al. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 2020;53:5455–516. https://doi.org/10.1007/s10462-020-09825-6.
[68] He K, Zhang X, Ren S, , Deep Residual Learning for Image Recognition. Proceedings of the IEEE conference on computer vision and pattern recognition;2016ap.770-778. ⟨https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html⟩.
[69] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14. Springer International Publishing; 2016. p. 630–45. https://doi.org/10.1007/978-3-319-46493-0_38.
[70] Xie S., Girshick R., Dollar P., et al. (2017) Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1492-1500). ⟨https://openaccess.thecvf.com/content_cvpr_2017/papers/Xie_Aggregated_Residual_Transformations_CVPR_2017_paper.pdf⟩.
[71] Zhang K, Tang B, Deng L, et al. A hybrid attention improved ResNet based fault diagnosis method of wind turbines gearbox. Measurement 2021;179:109491. https://doi.org/10.1016/j.measurement.2021.109491.
[72] Zhang H, Wu C, Zhang Z, , Resnest: Split-attention networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition;2022ap. 2736-2746⟨https://openaccess.thecvf.com/content/CVPR2022W/ECV/html/Zhang_ResNeSt_Split-Attention_Networks_CVPRW_2022_paper.html⟩.
[73] Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. Towards Data Sci 2017;6(12):310–6. ⟨https://www.ijeast.com/papers/310-316,Tesma412,IJEAST.pdf⟩.
[74] Shwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need. Inf Fusion 2022;81:84–90. https://doi.org/10.1016/j.inffus.2021.11.011.
[75] Zhang H, Cheng X, Li Y, et al. Prediction of failure modes, strength, and deformation capacity of RC shear walls through machine learning. J Build Eng 2022;50:104145. https://doi.org/10.1016/j.jobe.2022.104145.
[76] Mangalathu S, Karthikeyan K, Feng DC, et al. Machine-learning interpretability techniques for seismic performance assessment of infrastructure systems. Eng Struct 2022;250:112883. https://doi.org/10.1016/j.engstruct.2021.112883.
[77] Todorov B, Billah AM. Machine learning driven seismic performance limit state identification for performance-based seismic design of bridge piers. Eng Struct 2022;255:113919. https://doi.org/10.1016/j.engstruct.2022.113919.
[78] Arik SÖ, Pfister T. Tabnet: attentive interpretable tabular learning. Proc AAAI Conf Artif Intell 2021;35(8):6679–87. https://doi.org/10.1609/aaai.v35i8.16826.
[79] Sarrafzadeh M, Elwood KJ, Dhakal RP, et al. Performance of reinforced concrete buildings in the 2016 Kumamoto earthquakes and seismic design in Japan. Bull NZ Soc Earthq Eng 2017;50(3):394–435. https://doi.org/10.5459/bnzsee.50.3.394-435.
[80] Paszke A, Gross S, Massa F, et al. Pytorch: an imperative style, high-performance deep learning library. Adv Neural Inf Process Syst 2019;32. ⟨https://github.com/pytorch/pytorch⟩.
[81] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv prepr. arXiv 2017. https://doi.org/10.48550/arXiv.1412.6980. 1412.6980.
[82] Akiba T, Sano S, Yanase T, et al. (2019) Optuna: A next-generation hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD International

Conference on Knowledge Discovery & Data Mining, 2623-2631. https://doi.org/10.1145/3292500.3330701.
[83] FEMA, Multi-hazard Loss Estimation Methodology: Earthquake Model, HAZUS Technical Manual), 2003.
[84] Pagni CA, Lowes LN. Fragility functions for older reinforced concrete beam-column joints. Earthq Spectra 2006;22(1):215–38. https://doi.org/10.1193/1.2163365.
[85] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017:4765–74. ⟨https://arxiv.org/abs/1705.07874⟩.
[86] Ribeiro MT, Singh S, Guestrin C, Why should I trust you?": explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining;2016p. 1135-1144https://doi.org/10.1145/2939672.2939778.
[87] Sundararajan M, Taly A, Yan Q, Axiomatic attribution for deep networks. Proceedings of the 34th International Conference on Machine Learning; 2017p.3319-3328⟨https://arxiv.org/abs/1703.01365⟩.
[88] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29(5):1189–232. ⟨https://www.jstor.org/stable/2699986⟩.
[89] Borisov V, Leemann T, Seßler K, et al. Deep neural networks and tabular data: a survey. IEEE Trans Neural Netw Learn Syst 2022;35(6):7499–519. https://doi.org/10.1109/TNNLS.2022.3229161.