Proceedings of the
Annual Stability Conference
Structural Stability Research Council
Louisville, Kentucky, April 2-4, 2025

Mechanics-informed data-driven prediction model of steel column strength

Zhidong Zhang[1] Kai Chen[2] Cem Bartu Cevik[3] Cristopher D. Moen[4]

Abstract

The main objective of this study is to investigate the ability and application of mechanics-informed neural network models to predict steel column strength. Steel column strength prediction involves challenges associated with stability issues, buckling deformation, and initial geometric imperfections, and data-driven models informed by mechanics demonstrate significant potential for better accuracy and efficiency. A multilayer perceptron neural network is trained and tested. The training data is generated by sampling from the classical ordinary differential equation for a simply-supported steel column with an initial sweep imperfection combined with a first yield stress interaction failure criterion consistent with the Perry-Robertson equation. Prediction accuracy is evaluated for different network configurations and hyperparameters with the goal of establishing a model protocol that can be generally followed for mechanics-informed data models. Finally, governing equations are introduced to the loss function to further inform the neural network of more mechanics-based information. The performance of the mechanics-informed artificial neural network is discussed. This research develops high-quality data-driven models with mechanics-informed protocol enabled, which can further advance state-of-the-art steel column design and research.

1. Introduction

The application of Machine Learning (ML) methods to solve various structural engineering problems has become an area of active research in Civil Engineering in recent years, such as structural health monitoring (Stephens et al. 1994, Wu et al. 1992, Huang et al. 2019, and Sun et al. 2020), structural topology optimization (Berke et al. 1993, Kaveh 2024, and Shin et al. 2023), structural design and analysis (Hajela and Berke 1991, Salehi and Burgueño 2018, Guo et al. 2021, and Thai 2022), and other applications (Kaveh et al. 2024). Recent advancements in the computing power for Artificial Intelligence (AI), such as the emergence of highly efficient graphics processing units (GPUs) and other new-generation chips, have been a key driver of the further progress for the

[1]Postdoctoral Research Associate, Dept. of Civil and Environmental Engineering, University of Virginia, <zz9km@virginia.edu>

[2]Graduate Research Assistant, Dept. of Applied Mathematics and Statistics, Johns Hopkins University, <kchen137@jhu.edu>

[3]CTO, Datatoolpack, LLC, <bartu@datatoolpack.com>

[4]President, RunToSolve, LLC, <cris.moen@runtosolve.com>

application and research in machine learning-based numerical solutions, which empowers more efficient and accurate prediction from ML models.

Current research in the application of structural design and analysis predominantly employs ML regression algorithms, relying on large datasets (mainly numerically generated dataset through validated models) to train various ML models, including random forest, support vector machine, and artificial neural networks (ANNs). However, in practical structural engineering, data scarcity and high acquisition costs (no matter through experimental studies or simulation) hinder the widespread adoption of ML techniques in structural design and analysis. As a result, a novel machine learning technique, termed Physics-Informed Neural Networks (PINNs), has been recently proposed by Raissi et al. (2019). This unsupervised learning approach demonstrates significant potential for addressing intricate mechanical challenges through utilizing the physical information to guide the learning process and lowering the data-intensive demands. By incorporating the knowledge of physical laws during training, PINNs act as a form of regularization, restricting the search space of possible solutions and thereby improving the model's ability to generalize from limited data. (Katsikis et al. 2022, Yuan et al. 2022, and Zhang et al. 2020). It excels in scenarios where governing equations are known, eliminating the need for extensive pre-defined datasets and the solution process can be remarkably fast when the PINN is pre-trained and applicable to a transfer learning technique. The PINN application research of Chen et al. (2023) investigated its applicability of second-order analysis of beam-columns. The successful application of PINNs to static beam problems by Katsikis et al. (2022), where they accurately predicted linear behavior, supports their feasibility as a potential next-generation alternative method for structural analysis and design.

The object of this research is to provide a mechanics-informed data-driven modeling protocol to offer a promising approach to enhance accuracy, efficiency, and adaptability in the prediction of column strength through ML-powered engines. This study employs a multilayer perceptron neural network, trained and tested using data generated by sampling from the classical ordinary differential equation governing a simply supported steel column with an initial sweep imperfection. The failure criterion is based on a first yield stress interaction, consistent with the Perry-Robertson equation. The research investigates the impact of different network configurations and hyperparameters on model accuracy to establish a robust and generalizable protocol for mechanics-informed data-driven models. Furthermore, the study incorporates the Perry-Robertson equation as the governing equation into the loss function, enhancing the neural network's understanding of underlying physics. The decent performance of this mechanics-informed artificial neural network demonstrates its potential to advance the state-of-the-art in steel column design and research.

2. Dataset Generation and Processing

The data generation relies on the Perry-Robertson formula, as presented in Equation 1. To increase the data veracity, a random noise with mean value of 0 and standard deviation of 1 is added to the strength value P_n.

$$P_n = 0.5 \cdot \left(1 + \frac{\sigma_e}{\sigma_y} \cdot (\eta + 1) - \sqrt{\left(1 + \frac{\sigma_e}{\sigma_y} \cdot (\eta + 1)\right)^2 - 4 \cdot \frac{\sigma_e}{\sigma_y}}\right) \cdot f_y \cdot A \tag{1}$$

where:

- P_n is the axial load-bearing capacity of the column (kip).
- $\sigma_{cre} = \frac{\pi^2 EI}{AL^2}$ is the critical buckling axial stress defined by Euler's critical load equation (kip).
- σ_y is the yield strength of the column material (ksi).
- $\eta = \frac{\delta_o c}{\frac{I}{A}}$ is a factor considering the initial imperfection.
- $\delta_o = \frac{L}{1000}$ is the initial imperfection or out-of-straightness of the column (inch).
- c is the distance from the centroid to the outermost fiber of the column's cross-section (inch).
- I is the moment of inertia of the column's cross-section (inch^4).
- A is the cross-sectional area of the column (inch^2).
- L is the effective length of the column (inch).
- E is the elastic modulus of the column material (ksi).

The Perry-Robertson equation is widely used in the design of structural elements subjected to axial loads to consider stability, buckling deformation, and initial geometric imperfections. The data generated has the size of 1,000,000 rows $\times$ 7 columns. The seven columns corresponds to A (cross-sectional area), I (moment of inertia), c (distance from centroid to the outermost fiber), L (member length), f_y (yield strength), δ_o (the initial imperfection or out-of-straightness of the column), and P_n (axial compression strength). The data set can be downloaded from `Kaggle`, an open-source data science online community.

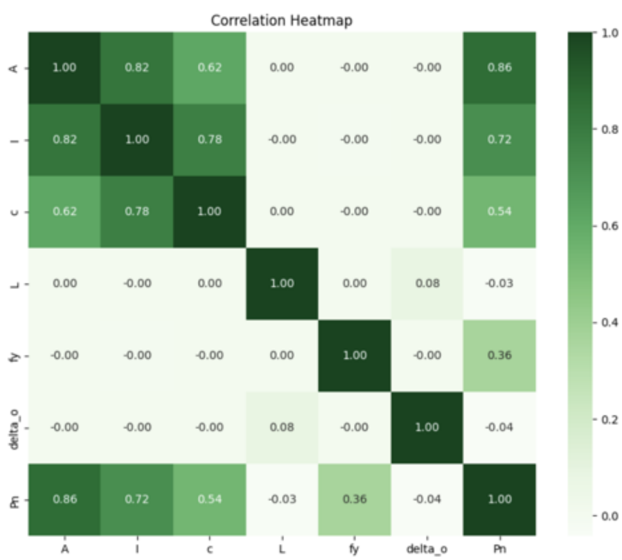

Figure 1: Correlation heatmap of the dataset

The correlation between the input and output features are evaluated through the pearson correlation coefficient (PCC), which implies a measure of the linear correlation between two sets of data. PCC is defined as the ratio between the covariance of two variables and the product of their standard deviations. To better present the PCC of the dataset, a heatmap is generated, as presented in Figure 1. Strong correlations with axial strength P_n exist for A (cross-sectional area), I (moment of inertia), and c (distance from centroid to the outermost fiber), featuring PCC values as 0.86, 0.72, and 0.54 respectively. f_y (yield strength) demonstrates moderate positive correlation with P_n (0.36). In terms of the correlations between input features, strong positive correlations are found between A and I, I and c, and A and c, which is reasonable in terms of section properties.

The generated dataset is stored in a comma-separated values (CSV) file and read in through a DataFrame format. To pre-process the data before inputting them into the neural networks, it is necessary to normalize the data to prevent dominance of certain features, improve gradient descent (convergence), and increase regularization performance. The input features of the dataset was scaled using the 'z-score normalization' method. Another pre-processing operation herein is the logarithmic normalization, which is applied to the output feature column axial strength P_n because it will always be a positive value. There are two considerations for applying the logarithmic normalization: (1) it can reduce and stabilize the variance and compress the range of P_n values, making the neural network model learn patterns within the dataset better and more efficiently; and (2) it ensures P_n is more aligned with a Gaussian distribution, which handles the skewness well and further improves model performance. Subsequently, the data was divided into a training set (85%) and a test set (15%) and input into the neural network models.

3. PINN Model Formulation

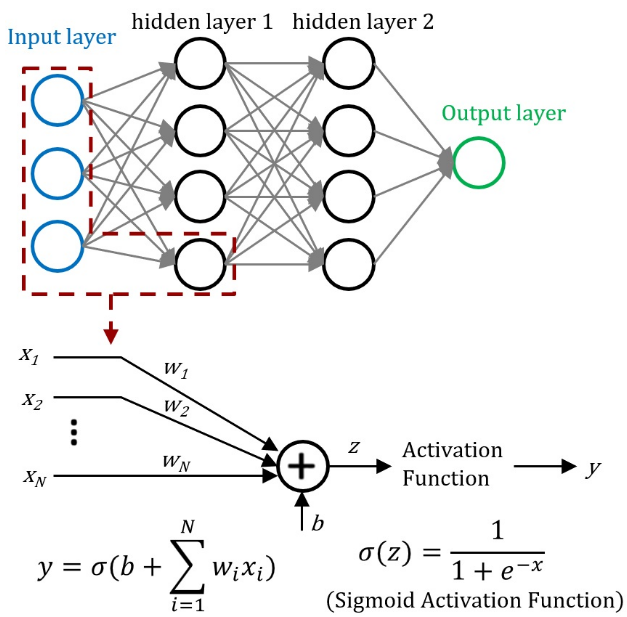

Figure 2: The architecture of a multilayer perceptron neural network

As shown in Figure 2, a multilayer perceptron neural network model is a basic architecture for an artificial neural network, which consists of multiple layers of neurons. There are input, output, and hidden layers in the main architecture frame. Weights (w_i) and biases (b) are defined at each neuron. Additionally, neuron values from the former layer are employed to generate the neuron value at the current layer using nonlinear transformation functions. Within the transformation function, activation functions introduce nonlinearity to the network, enabling it to learn complex patterns. Overall, the model learns patterns or representations from the input data at hidden layers.

For the training process of the multilayer perceptron neural network model, there are three main steps: forward propagation, loss evaluation, and backpropagation. Forward propagation involves sequentially feeding the input through the network's layers. Each layer processes the input by applying its unique set of weights, biases, and activation function, generating an output that serves as the input for the subsequent layer. The loss evaluation refers to the calculation of the loss function, which quantifies the discrepancy between the network's predictions and the ground truth. The backpropagation process involves iteratively refining the model's parameters (weights and biases) by minimizing the loss function. This optimization process is typically achieved through gradient descent or its variants. The training mainly happens at the backpropagation step when the weights and biases are updated and saved based on the optimization algorithms.

Before the development of a PINN model, we started with constructing a multilayer perceptron neural network model, and the accuracy of different network configurations and hyperparameters were evaluated. After the multilayer perceptron neural network model with best performance was determined, the loss function was then modified to include the mechanics-based information. The best performance neural network configuration demonstrates five layers, 2048 neurons in the hidden layers, Relu activation function ($max(0, x)$). Other parameters for the best neural network configuration include batch size, dropout, and batch normalization. Batch size implies the number of training examples utilized in one iteration, where larger batch sizes demonstrate more stable gradients, faster computation, but may generalize less; and smaller batch sizes feature noisier gradients, slower computation, but may generalize better. The batch size in the best model is set to be 256 demonstrating stable gradient and computational efficient. The dropout functionality is a regularization technique, where randomly selected neurons are ignored during training. This dropout functionality can force the network to learn redundant representations, reduces interdependence among neurons, and make the model more robust and less likely to overfit the dataset. The batch normalization technique can improve the speed, performance, and stability of neural networks, which normalizes the activations of each layer and reduces internal covariate shift. The model with best performance adopts both dropout and batch normalization modules.

As discussed in the Introduction section, a PINN demonstrates advantages including better accuracy and requiring a smaller data volume. As presented in Equation 2, the main difference between a PINN and a traditional multilayer perceptron neural network lies in the definition of loss function. The loss function for a PINN adds one more physical loss term in addition to the tradition loss term. The mean squared error (MSE) is adopted for the traditional loss function term. The Physical loss term herein refers to the difference between predicted value based on the Perry-Robertson equation and the ground truth value. The adjustment factor α for the physical loss term is taken as 1 herein.

$$PINN\ Loss = Traditional\ Loss + \alpha * Physical\ Loss \tag{2}$$

As shown in Figure 3, after modifying the loss function based on Equation 2 of the configuration of the multilayer perceptron neural network with best performance, the loss function value of the training and validation process of the model through 400 epoches exhibits the overall convergence trend of the model.

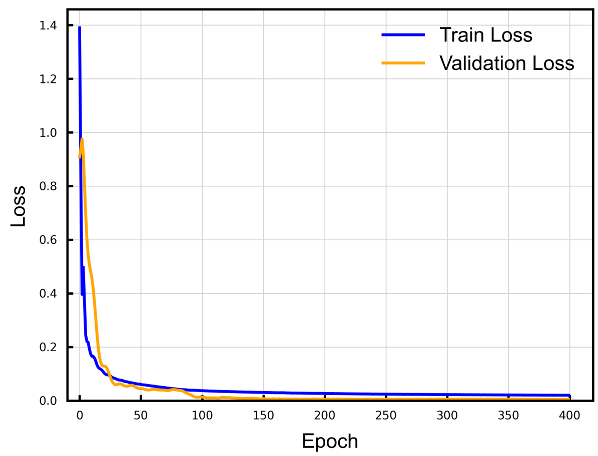

Figure 3: Train and validation loss curve

4. Results and Discussion

The preliminary results of the trained model's prediction is presented in Figure 4. The diagonal line of the figure implies that the horizontal axis value (the compressive strength P_n) and vertical axis value (predicted compressive strength P_n) are identical. The data points aligns with the diagonal line, although there are still data points demonstrating variance from the line. This shows that the PINN model learns the pattern behind the training dataset well.

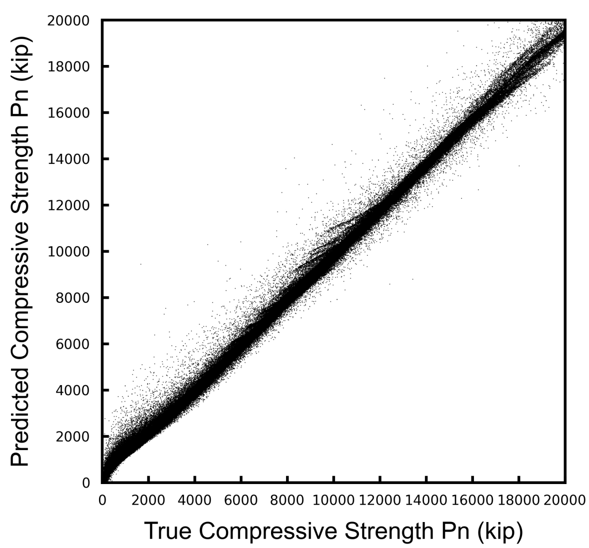

Figure 4: The comparision between prediction and ground truth values for P_n (kips)

To further explore the prediction performance of the trained PINN model, the prediction-to-ground truth values are plotted against the column slenderness (λ), as depicted in Figure 5. This shows that the overall prediction accuracy is acceptable along various scenarios with a wide range of column slenderness, but there still exists prediction error to be improved.

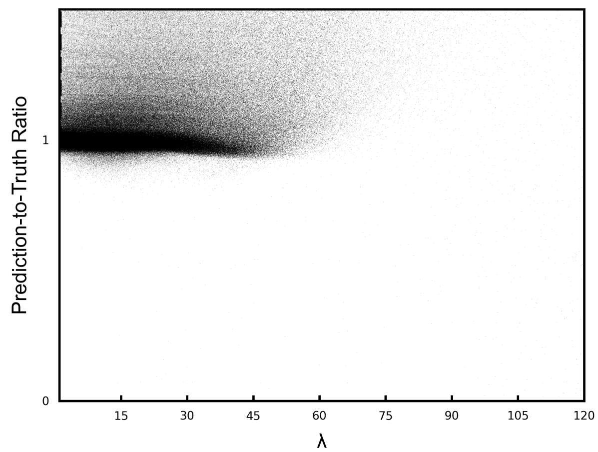

Figure 5: The distribution of prediction-to-truth ratio over column slenderness

5. Next Steps

The research reported in this paper is a preliminary study on the application of PINN for predicting steel column strength. Further fine tuning on the PINN model to improve its performance and accuracy is necessary. The next main research work is to compare the performance of PINN and traditional neural network to validate the PINN's performance on training efficiency and accuracy of prediction. At the same time, it is also an interesting topic to incorporate more governing equations into the physcial loss function to better guide the learning process of the neural network. Last but not least, further study on the appropriate adjustment factor α value to have better model performance is expected.

6. Conclusion

A multilayer perceptron neural network informed by mechanics is trained and tested in this research to predict the strength of steel columns. The dataset was generated by sampling from the classical ordinary differential equation consistent with the Perry-Robertson equation. Additionally, the Perry-Robertson equation was introduced to the loss function to further enhance the neural network's understanding of underlying physics. This mechanics-informed artificial neural network demonstrates decent performance in terms of column strength prediction. The reported work is still ongoing and more fine-tuning and comparison work will be completed in the near future. The mechanics-informed data-driven prediction model framework in this study exhibits great potential to advance the state-of-the-art in steel column design and research.

References

Berke, L., Patnaik, S., and Murthy, P. (1993). "Optimum design of aerospace structural components using neural networks". *Computers & structures* 48(6), 1001–1010.

Chen, L., Zhang, H.-Y., Liu, S.-W., and Chan, S.-L. (2023). "Second-order analysis of beam-columns by machine learning-based structural analysis through physics-informed neural networks". *Advanced Steel Construction* 19(4), 411–420.

Guo, K., Yang, Z., Yu, C.-H., and Buehler, M.J. (2021). "Artificial intelligence and machine learning in design of mechanical materials". *Materials Horizons* 8(4), 1153–1172.

Hajela, P. and Berke, L. (1991). “Neurobiological computational models in structural analysis and design”. *Computers & Structures* 41(4), 657–667.

Huang, H. and Burton, H.V. (2019). “Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning”. *Journal of Building Engineering* 25, 100767.

Katsikis, D., Muradova, A.D., and Stavroulakis, G.E. (2022). “A gentle introduction to physics-informed neural networks, with applications in static rod and beam problems”. *Journal of Advances in Applied & Computational Mathematics* 9, 103–128.

Kaveh, A. (2024). “Applications of artificial neural networks and machine learning in civil engineering”. *Studies in computational intelligence* 1168, 472.

Raissi, M., Perdikaris, P., and Karniadakis, G.E. (2019). “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. *Journal of Computational physics* 378, 686–707.

Salehi, H. and Burgueño, R. (2018). “Emerging artificial intelligence methods in structural engineering”. *Engineering structures* 171, 170–189.

Shin, S., Shin, D., and Kang, N. (2023). “Topology optimization via machine learning and deep learning: A review”. *Journal of Computational Design and Engineering* 10(4), 1736–1766.

Stephens, J.E. and VanLuchene, R.D. (1994). “Integrated assessment of seismic damage in structures”. *Computer-Aided Civil and Infrastructure Engineering* 9(2), 119–128.

Sun, L., Shang, Z., Xia, Y., Bhowmick, S., and Nagarajaiah, S. (2020). “Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection”. *Journal of Structural Engineering* 146(5), 04020073.

Thai, H.-T. (2022). “Machine learning for structural engineering: A state-of-the-art review”. *Structures*. Vol. 38. Elsevier, 448–491.

Wu, X., Ghaboussi, J., and Garrett Jr, J. (1992). “Use of neural networks in detection of structural damage”. *Computers & structures* 42(4), 649–659.

Yuan, L., Ni, Y.-Q., Deng, X.-Y., and Hao, S. (2022). “A-PINN: Auxiliary physics informed neural networks for forward and inverse problems of nonlinear integro-differential equations”. *Journal of Computational Physics* 462, 111260.

Zhang, R., Liu, Y., and Sun, H. (2020). “Physics-informed multi-LSTM networks for metamodeling of nonlinear structures”. *Computer Methods in Applied Mechanics and Engineering* 369, 113226.