

International Journal Research Publication Analysis

Page: 01-09

COMPARATIVE RESULTS OF MACHINE LEARNING MODELS FOR SEISMIC RESPONSE AND DAMAGE ASSESSMENT OF REINFORCED CONCRETE BUILDINGS

Ayush Raikwar^{*1}, Aditya Lakhera², Dr. Mayur Singi³

¹MTech Scholar, Department of Civil Engineering, Shri Krishna University, Chhatarpur,
India.

²Guide, Assistant Professor, Department of Civil Engineering, Shri Krishna University,
Chhatarpur, India.

³Co-Guide, Assistant Professor, Department of Civil Engineering, Shri Krishna University,
Chhatarpur, India

Article Received: 26 August 2026

Article Revised: 14 September 2026

Published on: 04 October 2026

*Corresponding Author: Ayush Raikwar

MTech Scholar, Department of Civil Engineering, Shri Krishna University, Chhatarpur,
India.

DOI: <https://doi-doi.org/101555/ijrpa.3007>

ABSTRACT

This paper presents a results-focused comparative assessment of machine learning approaches for seismic response and damage prediction of reinforced concrete (RC) buildings. The analysis is based on published benchmark results documented in the thesis source and is explicitly separated from original student-generated numerical results. For maximum interstory drift ratio (IDR_{max}), the reported benchmark dataset compares Random Forest, Bagging, Extra Trees, HistGradientBoosting, ANN, XGBoost, KNN and conventional regression models using R^2 , RMSE and MAE. Random Forest reports $R^2=0.901$, while Bagging reports the lowest RMSE and MAE among the highlighted high-performing models. Additional published results show prediction accuracies of approximately 90.2–96.3% for selected six-story RC configurations and 93.4–95.3% for eight-story configurations. For real earthquake damage validation, Random Forest reported 74.62% testing accuracy compared with 62.68% for KNN. The paper interprets these results from an engineering perspective and identifies the implications of nonlinear feature interactions, independent validation and model selection. All numerical values are labelled as literature benchmarks; they are not claimed as original Ayush Raikwar experimental or numerical results.

KEYWORDS: IDRmax; Random Forest; XGBoost; ANN; seismic damage; RC buildings; RMSE; MAE; machine learning.

1. INTRODUCTION

Seismic response prediction requires consideration of both structural characteristics and ground-motion variability. The thesis source organizes the problem around maximum interstory drift ratio, damage state and rapid prediction. Because the available thesis chapter explicitly states that its reported numerical values are benchmark results from previous studies, this paper treats them as a comparative results dataset rather than as newly generated ETABS/OpenSees results. This distinction is important for academic integrity and allows the results to be used for discussion without misrepresenting the source.

2. Benchmark Dataset and Evaluation Measures

The principal benchmark for continuous response prediction is the study by Kazemi et al., which used 92,400 data points obtained from IDA of 165 RC moment-resisting frames. The reported feature set included structural descriptors and $S_a(T_1)$, with IDRmax as the principal response quantity. [1] The comparison includes R^2 , RMSE and MAE. Higher R^2 indicates stronger agreement between predicted and reference response, whereas lower RMSE and MAE indicate smaller prediction error. A second benchmark from Bhatta and Dang addresses categorical damage prediction using simulation data and real damage observations associated with the 2015 Nepal earthquake. [2]

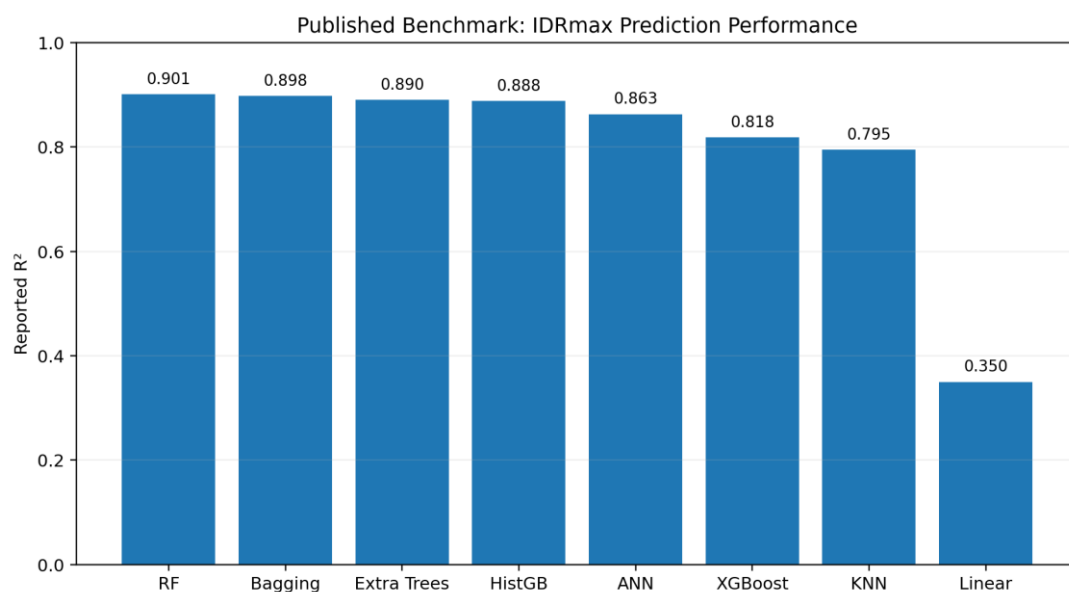


Fig. 1. Reported R^2 values for selected ML models in the IDRmax benchmark.

3. IDRmax Prediction Results.

Model	R ²	RMSE	MAE
RF	0.901	1.85	1.35
Bagging	0.898	1.44	1.04
Extra Trees	0.890	1.50	1.10
HistGB	0.888	1.51	1.07
ANN	0.863	1.79	1.35
XGBoost	0.818	1.61	1.24
KNN	0.795	1.78	1.17
Linear	0.350	2.60	2.05

The reported R² comparison shows that Random Forest, Bagging, Extra Trees and HistGradientBoosting form a high-performing group in the benchmark. Random Forest has the highest reported R² at 0.901, followed closely by Bagging at 0.898 and Extra Trees at 0.890. HistGradientBoosting reaches 0.888, while ANN and XGBoost report 0.863 and 0.818, respectively. KNN reaches 0.795. The conventional Linear Regression benchmark is much lower at 0.350, illustrating that a simple linear mapping is insufficient for the nonlinear relationship represented in the dataset. [1]

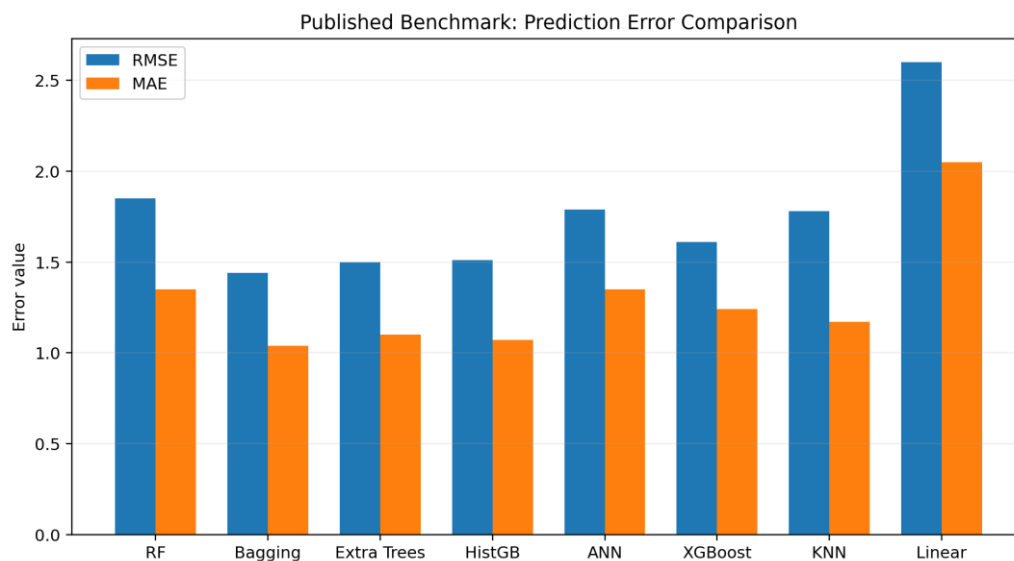


Fig. 2. Reported RMSE and MAE values for the same benchmark models.

4. Error-Based Interpretation

R² alone does not determine the most useful model. The reported Bagging model has RMSE=1.44 and MAE=1.04, both lower than the corresponding Random Forest values of 1.85 and 1.35. Extra Trees and HistGradientBoosting also show comparatively small errors. This illustrates an important engineering point: model selection should consider both correlation and absolute prediction error. A model can have the highest R² while another

model produces smaller average deviations. Accordingly, a final thesis implementation should report all three measures on the same independent test set rather than selecting a model from a single score.

5. Generalization Across Building Configurations

Published results reported for building configurations excluded from training provide an additional test of generalization. For selected six-storey RC frames with one to five bays, reported prediction accuracies were approximately 90.2%, 93.5%, 94.0%, 95.4% and 96.3%. For selected eight-storey configurations, reported accuracies were approximately 93.4–95.3%. [1] These results indicate that ML can retain useful predictive capability when the model is evaluated on structural configurations not directly used during training, provided the new cases remain within the domain represented by the overall database.

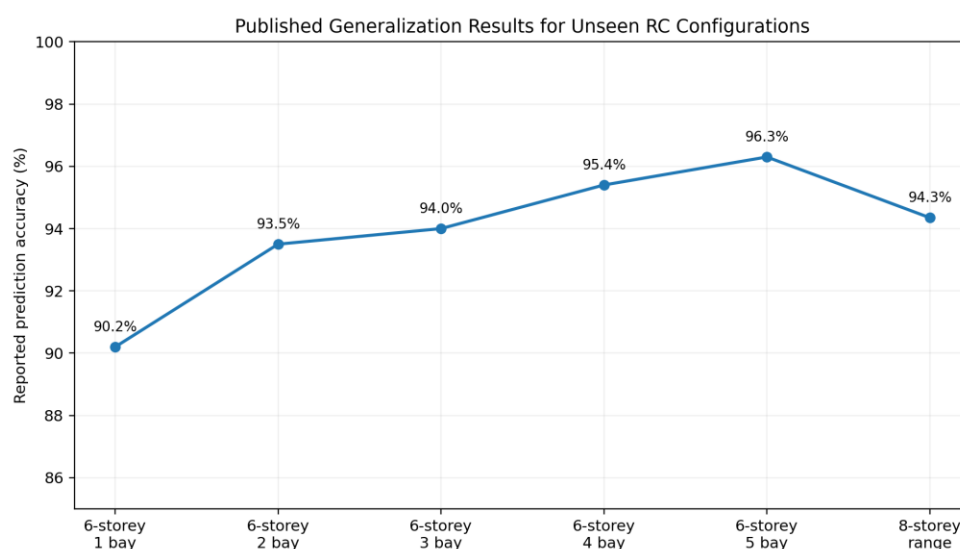


Fig. 3. Published prediction accuracies reported for selected unseen RC building configurations.

6. Seismic Damage-State Results

Bhatta and Dang provide a complementary classification benchmark. Their study considered 35 RC buildings, 11 earthquake waves and ML models including KNN, Random Forest, Decision Tree, SVM and ANN. The authors also used observed damage information from the 2015 Nepal earthquake for testing. Random Forest reported a training accuracy of 99.83% and a testing accuracy of 74.62%, whereas KNN reported 62.68% testing accuracy. [2] The difference between training and testing performance is itself a result: high training accuracy

cannot be treated as evidence of generalization. Real or externally generated observations provide a more demanding evaluation.

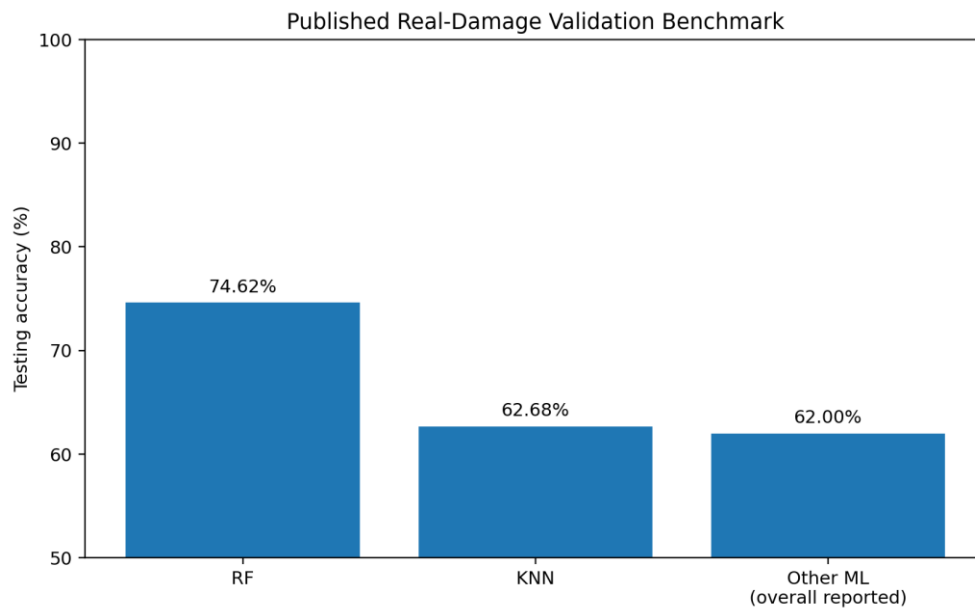


Fig. 4. Published testing-accuracy benchmark for real-damage validation.

7. Comparison with Recent Deep-Learning Results

More recent work extends the problem to cumulative mainshock-aftershock damage, post-earthquake damage detection, and regional seismic risk assessment using machine learning. [8] [25]

8. ENGINEERING DISCUSSION

The benchmark results indicate that nonlinear ensemble methods are well suited to the RC seismic-response problem because the response depends on interacting variables rather than one dominant linear relationship. The importance of stories, weight, number of bays, fundamental period and spectral acceleration is consistent with basic structural mechanics. However, the results also demonstrate why external validation is necessary. A model can perform extremely well on a simulated training set and still lose accuracy when applied to real earthquake damage. For a thesis-level implementation, the final result chapter should therefore include an independent building-level split, record-level separation, residual plots, actual-versus-predicted plots and feature-importance analysis. If the original ETABS/OpenSees/IDA database is later generated, these literature benchmarks can be retained only as comparison values.

The final original-results paper should report: (i) number of structural models, (ii) number and source of earthquake records, (iii) intensity scaling procedure, (iv) response extraction method, (v) train/validation/test split, (vi) hyperparameters, (vii) R^2 , RMSE and MAE for every model, (viii) actual-versus-predicted plots, (ix) residual distribution, (x) feature importance and (xi) independent validation on a building or record family excluded from training. This format will allow the student's own results to be distinguished clearly from published benchmark values.

9. Detailed Interpretation of the Benchmark Table

The benchmark table also demonstrates that different performance indicators can lead to different impressions of model quality. Random Forest has the highest reported R^2 among the listed models, but Bagging has lower RMSE and MAE. This is not a contradiction: R^2 measures the proportion of variance represented by the model, whereas RMSE and MAE describe the magnitude of prediction error. In engineering applications, a model should therefore be evaluated using a group of indicators and not selected only because one score is numerically highest.

The difference between ensemble methods and simple linear regression is also informative. Linear regression assumes that the target can be approximated by a weighted combination of the input variables. Seismic response of RC buildings does not generally follow such a simple relationship because stiffness, mass, period, intensity and nonlinear deformation interact. Tree ensembles can divide the feature space into regions and represent changing relationships across those regions. This provides a practical explanation for their stronger benchmark performance.

At high deformation levels, however, prediction becomes more challenging. Yielding, stiffness degradation, strength deterioration and possible instability can introduce patterns that are less represented in the training data. A model may therefore show strong average performance while producing larger residuals near severe response states. For this reason, the final original-results study should inspect errors separately for low, moderate and high drift ranges.

Generalization should also be tested structurally rather than only statistically. If observations from the same building appear in both training and test subsets, the model may learn building-specific patterns. A stronger design is to hold out complete building configurations or groups of earthquake records. The reported studies that evaluate unseen configurations provide useful examples for designing such a test.

Finally, the benchmark values should be used as comparison points rather than as substitute results. Once the actual Ayush Raikwar thesis dataset is available, the same metrics can be calculated using the same definitions and placed beside the published values. This will create a transparent comparison between the student's model and earlier studies without mixing datasets.

10. CONCLUSION

The comparative results demonstrate that machine learning can provide useful approximations of nonlinear seismic response and damage state when the training database is representative. In the published IDRmax benchmark, Random Forest has the highest reported R^2 , while Bagging and related ensemble models show low error values. Real-damage validation also shows that testing performance can be substantially lower than training performance, reinforcing the need for independent validation. The present paper should therefore be read as a literature-based results comparison. It does not claim that the reported numerical values were generated by Ayush Raikwar. The appropriate next step for an original result paper is to replace the benchmark tables with the student's actual structural-analysis and ML outputs once those are available.

REFERENCES

1. Kazemi, F., Asgarkhani, N., & Jankowski, R. (2023). Machine learning-based seismic response and performance assessment of reinforced concrete buildings. *Archives of Civil and Mechanical Engineering*, 23, 94. <https://doi.org/10.1007/s43452-023-00631-9>.
2. Bhatta, S., & Dang, J. (2023). Seismic damage prediction of RC buildings using machine learning. *Earthquake Engineering & Structural Dynamics*, 52(11), 3504–3527. <https://doi.org/10.1002/eqe.3907>.
3. Wen, W., Zhang, C., & Zhai, C. (2022). Rapid seismic response prediction of RC frames based on deep learning and limited building information. *Engineering Structures*, 267, 114638. <https://doi.org/10.1016/j.engstruct.2022.114638>.
4. Luo, H., & Paal, S. G. (2022). Artificial intelligence-enhanced seismic response prediction of reinforced concrete frames. *Advanced Engineering Informatics*, 52, 101568. <https://doi.org/10.1016/j.aei.2022.101568>.
5. Zhang, H., Cheng, X., Li, Y., He, D., & Du, X. (2023). Rapid seismic damage state assessment of RC frames using machine learning methods. *Journal of Building Engineering*, 65, 105797. <https://doi.org/10.1016/j.jobbe.2022.105797>.

6. Gondaliya, K. M., Vasanwala, S. A., Desai, A. K., Amin, J. A., & Bhaiya, V. (2024). Machine learning-based approach for assessing the seismic vulnerability of reinforced concrete frame buildings. *Journal of Building Engineering*, 97, 110785. <https://doi.org/10.1016/j.jobe.2024.110785>.
7. Xu, W., Zhao, Y., Yang, W., Yu, D., & Zhao, Y. (2024). Seismic fragility analysis of RC frame structures based on IDA analysis and machine learning. *Structures*, 65, 106774. <https://doi.org/10.1016/j.istruc.2024.106774>.
8. Gong, M., Liu, B., Wang, X., Zhou, B., Zhao, Y., & Jia, J. (2025). Damage assessment of reinforced concrete frame under mainshock-aftershock based on deep learning considering pre-earthquake damage. *Journal of Building Engineering*, 100, 111729. <https://doi.org/10.1016/j.jobe.2024.111729>.
9. Liu, B., Gong, M., Wang, X., Zhou, B., & Jia, J. (2026). Real-time seismic damage assessment framework of RC frames using deep learning and monitoring data. *Journal of Building Engineering*, 123, 115834. <https://doi.org/10.1016/j.jobe.2026.115834>.
10. Deneko, E., Hysenlliu, M., & Dhoska, K. (2026). Benchmarking machine learning for probabilistic seismic demand prediction of reinforced concrete frames in intelligent performance-based design. *Journal of Intelligent Construction*. <https://doi.org/10.26599/JIC.2026.9180136>.
11. Amini Pishro, A. (2026). Machine learning applications in reinforced concrete structures: A comprehensive review of methods, challenges, and future directions. *Structures*, 89, 112103. <https://doi.org/10.1016/j.istruc.2026.112103>.
12. Abubakar, M., & Awaib, M. (2026). Seismic analysis and probabilistic risk assessment of reinforced concrete buildings: a review and integrated framework. *Discover Civil Engineering*, 3, 179. <https://doi.org/10.1007/s44290-026-00574-z>.
13. Chopra, A. K. (2017). *Dynamics of Structures: Theory and Applications to Earthquake Engineering* (5th ed.). Pearson.
14. FEMA. (2009). *Quantification of Building Seismic Performance Factors* (FEMA P-695). Federal Emergency Management Agency, Washington, DC.
15. ASCE/SEI. (2022). *Minimum Design Loads and Associated Criteria for Buildings and Other Structures* (ASCE/SEI 7-22). American Society of Civil Engineers.
16. ASCE/SEI. (2017). *Seismic Evaluation and Retrofit of Existing Buildings* (ASCE/SEI 41-17). American Society of Civil Engineers.
17. McKenna, F., Scott, M. H., & Fenves, G. L. (2010). Nonlinear finite-element analysis software architecture using object composition. *Computing in Civil Engineering*, ASCE.

18. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
<https://doi.org/10.1023/A:1010933404324>.
19. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
20. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
21. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer.
22. Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
23. Hwang, S.-H., Mangalathu, S., Shin, J., & Jeon, J.-S. (2021). Machine learning-based approaches for seismic demand and collapse of ductile reinforced concrete building frames. *Journal of Building Engineering*, 34, 101905.
<https://doi.org/10.1016/j.jobbe.2020.101905>.
24. Zou, D., Zhang, M., Bai, Z., Liu, T., Zhou, A., Wang, X., Cui, W., & Zhang, S. (2022). Multicategory damage detection and safety assessment of post-earthquake reinforced concrete structures using deep learning. *Computer-Aided Civil and Infrastructure Engineering*, 37, 1188–1204. <https://doi.org/10.1111/mice.12815>.
25. Kourehpaz, P., & Molina Hutt, C. (2022). Machine Learning for Enhanced Regional Seismic Risk Assessments. *Journal of Structural Engineering*, 148(9).
[https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0003421](https://doi.org/10.1061/(ASCE)ST.1943-541X.0003421).