

PROCEEDINGS

Interpretable Prediction for Seismic Failure Modes of Corroded RC Columns Using Class-Imbalanced Data and Cost-Sensitive Learning

Jiannan Jiang¹ and Bo Yu^{1,2,*}¹School of Civil Engineering and Architecture, Guangxi University, Nanning, 530004, China²Key Laboratory of Engineering Disaster Prevention and Structural Safety of China Ministry of Education, Guangxi Laboratory of Life-Cycle Safety for Land-Sea Corridor Engineering, Guangxi Key Laboratory of Disaster Prevention and Engineering Safety, Nanning, 530004, China

*Corresponding Author: Bo Yu. Email: gxuyubo@gxu.edu.cn

ABSTRACT

Reinforced concrete (RC) columns may experience various failure modes under chloride-induced corrosion and seismic loading, including flexural, shear, and flexure-shear failures. However, existing datasets often suffer from class imbalance, where flexural failure samples vastly outnumber shear and flexure-shear failures. This imbalance leads to low prediction accuracy for brittle failure modes in traditional machine learning models, limiting seismic performance assessment of corroded RC columns. To address this issue, this paper proposes an interpretable ensemble learning model for predicting seismic failure modes of corroded RC columns using imbalanced data and cost-sensitive learning. First, a multi-class imbalanced dataset containing 289 experimental cases was constructed. Then, cost-sensitive learning was integrated into the Light Gradient Boosting Machine framework with an improved loss function incorporating class weight coefficients to better capture minority failure classes. Subsequently, Bayesian optimization and K-fold cross-validation were applied for hyperparameter tuning, and the SHapley Additive exPlanations (SHAP) method was integrated to enhance interpretability. The results indicate that the proposed model achieved an overall accuracy of 91%, about 15% higher than traditional models. The F1-scores for flexure-shear and shear failures reached 81%, with improvements of 27% and 21%, respectively. SHAP analysis revealed that the shear span ratio is the most critical parameter influencing failure modes. Finally, the model's Kappa coefficient of 0.83 indicates strong consistency with actual classifications. Under different scenarios with varying degrees of class imbalance, the proposed model consistently outperformed other models, validating its effectiveness in handling imbalanced data and improving the identification of minority classes.

KEYWORDS

Corroded RC column; seismic failure mode; class imbalance; cost sensitive learning; SHAP interpretability

Funding Statement: National Natural Science Foundation of China, 52278162 and 62266005, B. Yu. Guangxi Key Research and Development Project, GKAB23026026 and GKAB23075207, B. Yu. Guangxi Science and Technology Major Project, GKAA23023034 and GKAA23073017, B. Yu.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.



Copyright © 2026 The Author(s). Published by Tech Science Press.

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.