Journal of Zhejiang University-SCIENCE A

Predicting permeability coefficients of earth-rock material using an improved generative adversarial network and explainable ensemble learning under small sample conditions

Chengyu YU, Hongling YU, Xiaofeng QU, Baoxi LIU, Liangsi XU, Xinyu LIU, Xiangyu CHEN

Cite this as: Chengyu YU, Hongling YU, Xiaofeng QU, Baoxi LIU, Liangsi XU, Xinyu LIU, Xiangyu CHEN, 2026. Predicting permeability coefficients of earth-rock material using an improved generative adversarial network and explainable ensemble learning under small sample conditions. *Journal of Zhejiang University-SCIENCE A*, 27(3):215-230. https://doi.org/10.1631/jzus.A2500127

The contributions of this research

Data Augmentation Technology Based on WCGAN

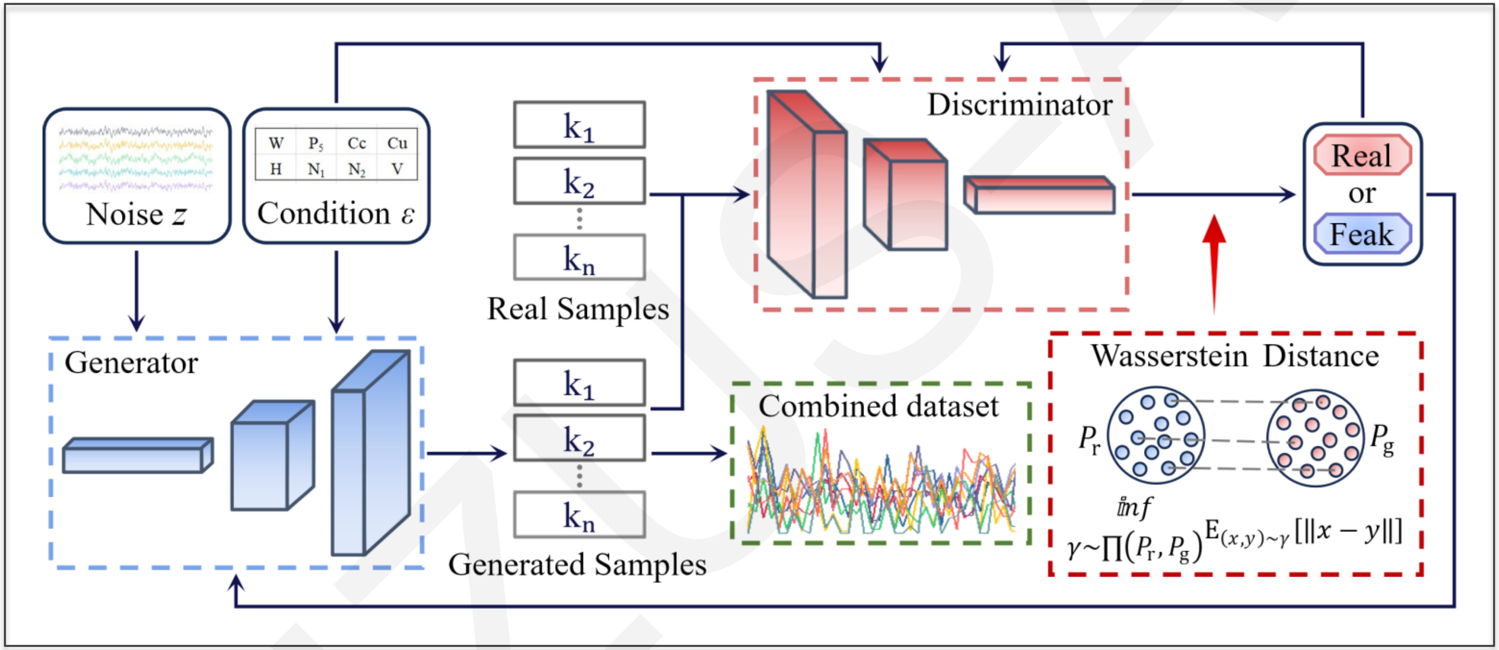

Fig. 1 Schematic diagram of the working principle of WCGAN.

Proposed a data augmentation method based on improved GAN (WCGAN), guiding sample generation through CGAN and introducing Wasserstein distance as the loss function for CGAN, which addresses small-sample problems to a certain extent.

The contributions of this research

Prediction Model Based on OOA-HL-LightGBM

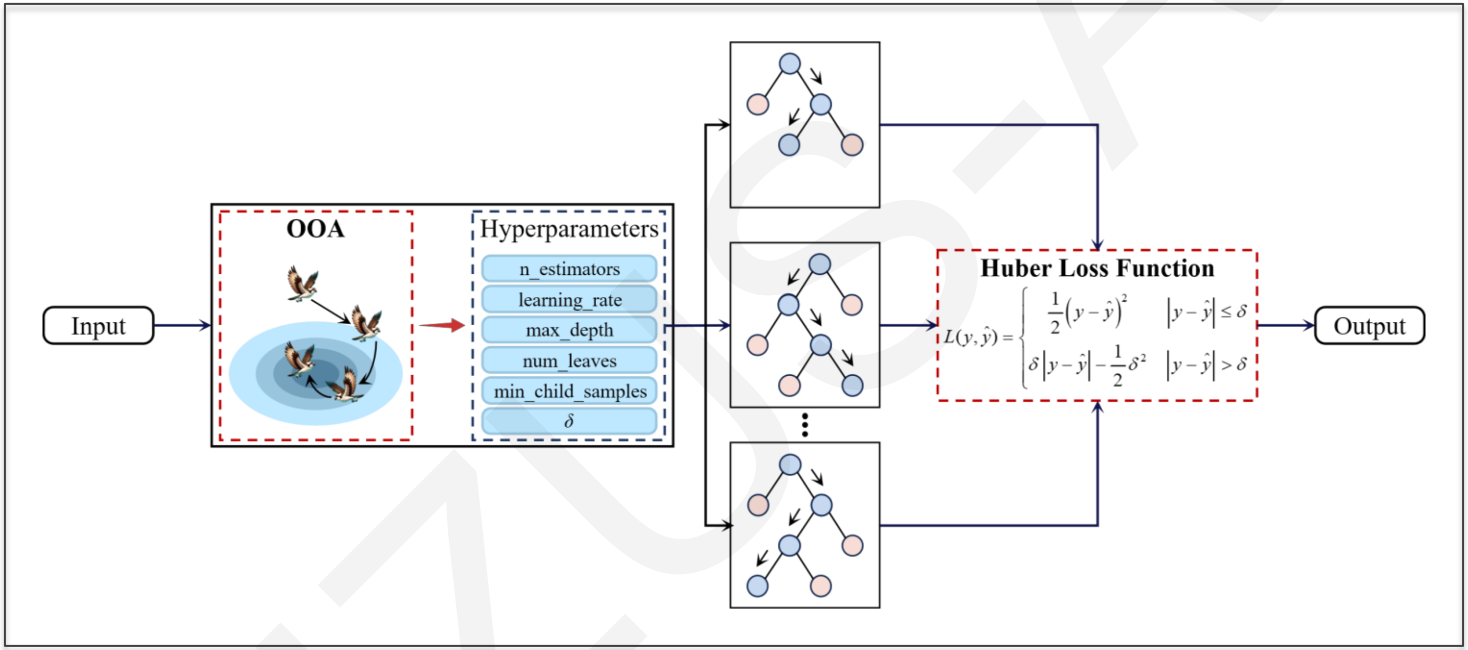

Fig. 2 Schematic diagram of the OOA-HL-LightGBM prediction model.

Constructed a prediction model based on improved LightGBM, incorporating Huber loss to mitigate the impact of outliers on prediction accuracy, and utilizing an optimization algorithm to optimize the hyperparameters, thereby achieving high-precision prediction of permeability coefficients.

The contributions of this research

Explainability Analysis Based on SHAP

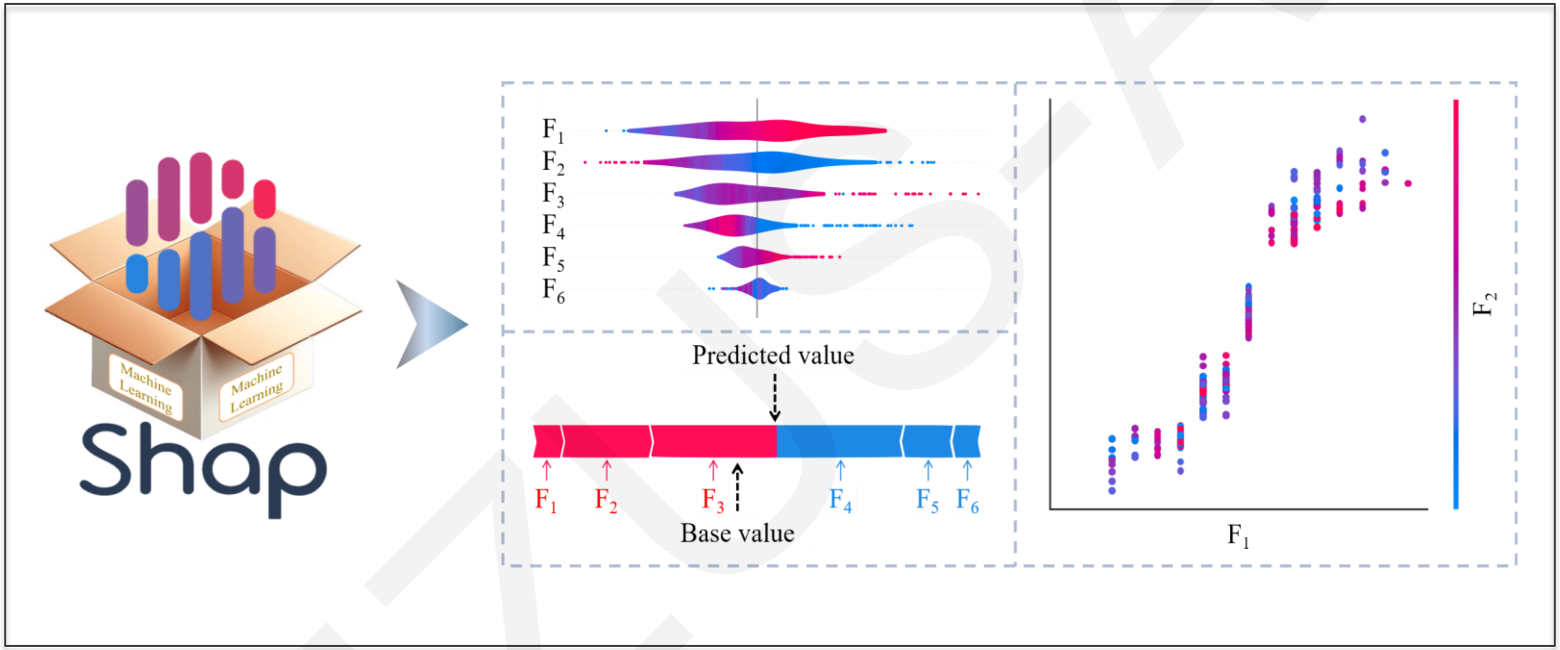

Fig. 3 Schematic diagram of the SHAP-based explainability mechanism.

Incorporated SHAP to provide explanations of the model from both global and local perspectives, identify the most critical features affecting prediction results, and analyze the specific roles of different features within the samples, thereby enhancing model explainability.

Main conclusions

01

The proposed WCGAN guides sample generation through the CGAN mechanism and Wasserstein distance, generating high-quality samples stably and effectively addressing the issue of small sample data. It demonstrates superior performance in model accuracy compared to GMM-KNN and VAE-based methods.

02

The OOA-HL-LightGBM model, which introduces Huber loss and utilizes OOA for hyperparameter optimization, achieves high-precision prediction of the permeability coefficient. Comparative analyses demonstrate that the improved approach significantly outperforms traditional algorithms in predictive performance.

03

The integration of SHAP attribution analysis with the prediction model facilitates the identification of the most critical features affecting prediction results and the analysis of the specific roles of different features within the samples, improving the credibility of prediction results for engineering applications.