Supplementary Materials

Data-driven design of high-abundance rare earth permanent magnets for low-carbon and resource-efficient production

Yuxuan Wu[1,3], Qianqian Wang[2,*], Qihan Zhang[3], Shen Zhao[3], Qiangfeng Li[3], Dehai Wu[4], Fang Wu[4], Hongdong Yu[1,3], Lu Wang[1,3,4,*]

[1]School of Rare Earths, University of Science and Technology of China, Hefei 230026, Anhui, China.

[2]School of Low-Altitude Technology and Engineering, Gannan College of Science and Technology, Ganzhou 341000, Jiangxi, China.

[3]Ganjiang Innovation Academy, Chinese Academy of Sciences, Ganzhou 341119, Jiangxi, China.

[4]Key Laboratory of Ionic Rare Earth Resources and Environment, Ministry of Natural Resources of the People's Republic of China, Ganzhou 341000, Jiangxi, China.

[*]**Correspondence to:** Prof. Qianqian Wang, School of Low-Altitude Technology and Engineering, Gannan College of Science and Technology, Ganzhou 341000, Jiangxi, China. E-mail: wangqianqian_cn@163.com; Prof. Lu Wang, Ganjiang Innovation Academy, Chinese Academy of Sciences, Ganzhou 341119, Jiangxi, China. E-mail: lwang@gia.cas.cn

Supplementary Table 1. Essential parameters and settings used in the MLP regression model

Parameter	Setting
Model	MLP regression
Hidden layer structure	(128, 64, 32)
Activation function	relu
Optimizer	Adam
Maximum iterations	3000
Random seed	42
Data scaling	Min–Max normalization
Early stopping	Enabled
Validation fraction	0.15
L2 regularization coefficient (α)	1.048×10^{-6}
Batch size	16
Learning rate	9.65×10^{-3}
Hyperparameter optimization method	Bayesian optimization
Bayesian optimization iterations	30
Optimized hyperparameters	α, learning rate, batch size
α search range	1×10^{-6} – 1×10^{-2}
Learning rate search range	1×10^{-4} – 1×10^{-2}
Batch size search range	16, 32, 64
Validation method	Repeated five-fold cross-validation
Number of folds	5
Number of repeats	3

Supplementary Table 2. Essential parameters and settings used in the NSGA-II model

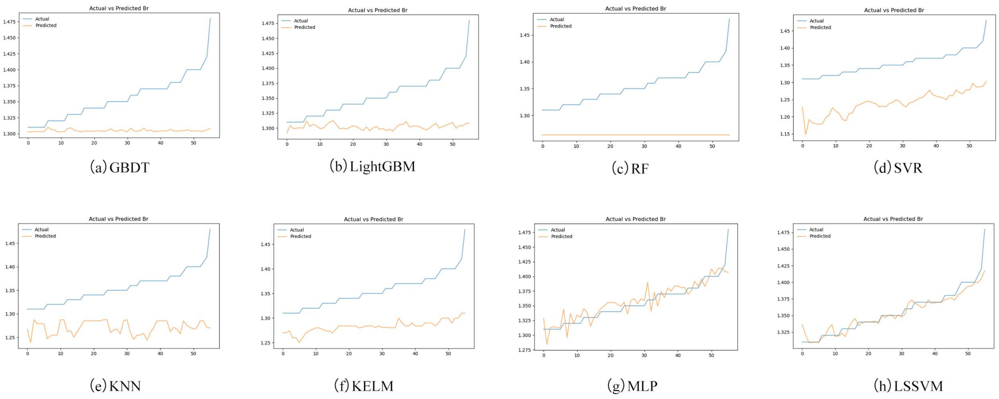
Parameter	Setting
Number of decision variables	10
Number of objectives	3
Lower bounds (varmin)	[0,0,0,0,0, 1,10,0,0,1]
Upper bounds (varmax)	[1,1,1,1,1,35,100,10,0.01,20]
Population size	npop = 700
Maximum iterations	maxit = 500
Crossover probability	pc = 0.7
Number of offspring (nc)	round(pc*npop/2)*2
Arithmetic factor	gamma = 0.1
Mutation probability	mu = 0.3
Step size	sigma = 0.05
First 5 vars perturbation	delta = 0.1*randn
Nonnegativity enforcement	max(y2,0)

Supplementary Table 3. Statistics of the datasets

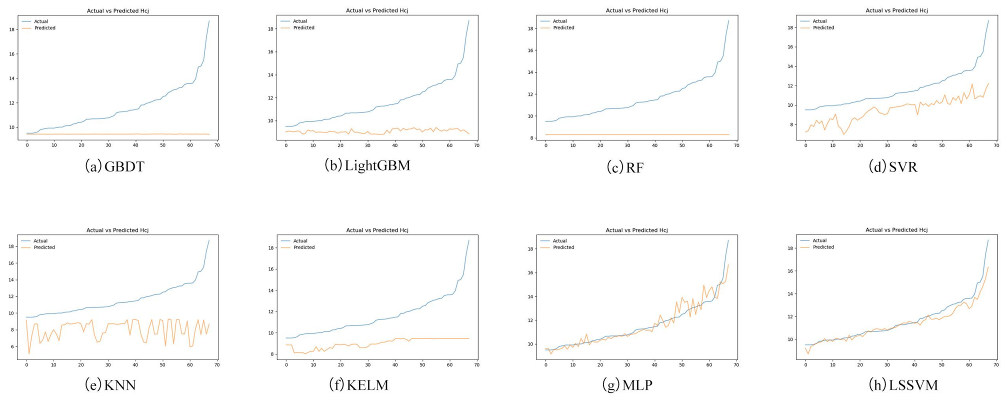
Variables		Min	Max	Mean	Median	Standard deviation
Input variables	P_{Ce}	0	1	0.12	0.08	0.15
	P_{La}	0	1	0.36	0.26	0.32
	P_{Y}	0	1	0.10	0	0.26
	P_{Pr}	0	1	0.04	0.05	0.11
	P_{Nd}	0	1	0.37	0.15	0.37
	REs	7	33.50	15.40	13.7	5.32
	P_{Fe}	60	91	78.17	80	5.08
	P_{Co}	0	17	0.24	0	1.75
	P_{TM}	0	8.50	0.28	0	0.89
	P_{B}	0.95	24	5.95	6.05	2.68
Output variables	Br	0.20	1.48	0.97	1.15	0.39
	Hcj	0.11	18.70	5.60	5.03	3.94
	(BH)max	0.19	48.10	16.51	7.56	15.91

Supplementary Table 4. Performance evaluation of the MLP models for predicting Br, H_{cj}, and $(BH)_{max}$ under different train-test splitting ratios

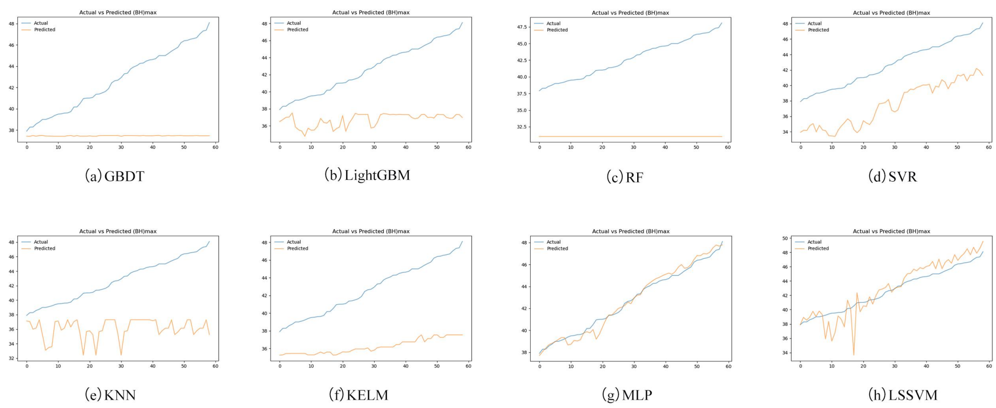
Target property	Split ratio (Train/Test)	Train R^2	Test R^2	Cross-validation R^2 (Mean ± SD)
	70/30	0.969	0.967	0.910 ± 0.065
Br	80/20	0.970	0.979	0.936 ± 0.027
	90/10	0.980	0.982	0.941 ± 0.034
	70/30	0.950	0.866	0.841 ± 0.086
Hcj	80/20	0.950	0.901	0.826 ± 0.068
	90/10	0.942	0.934	0.821 ± 0.074
	70/30	0.695	0.668	0.816 ± 0.112
$(BH)_{max}$	80/20	0.978	0.955	0.850 ± 0.091
	90/10	0.960	0.971	0.857 ± 0.082



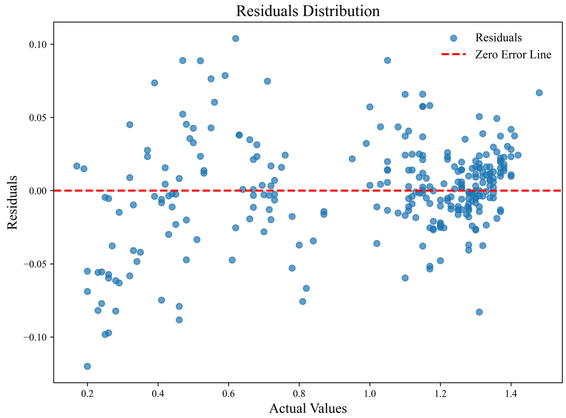
Supplementary Figure 1. Performance comparison of eight machine learning models for $(BH)_{max}$ prediction (actual vs. predicted values).



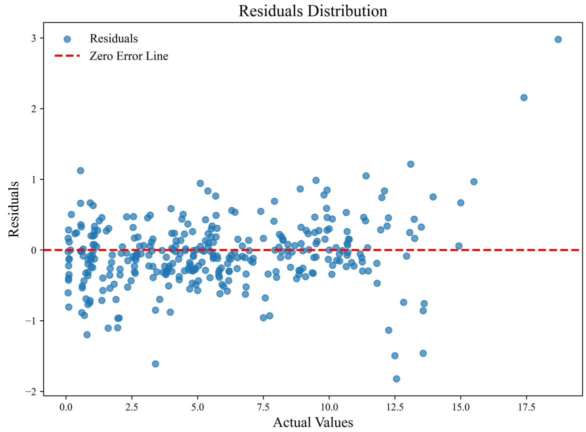
Supplementary Figure 2. Performance comparison of eight machine learning models for Br prediction (actual vs. predicted values).



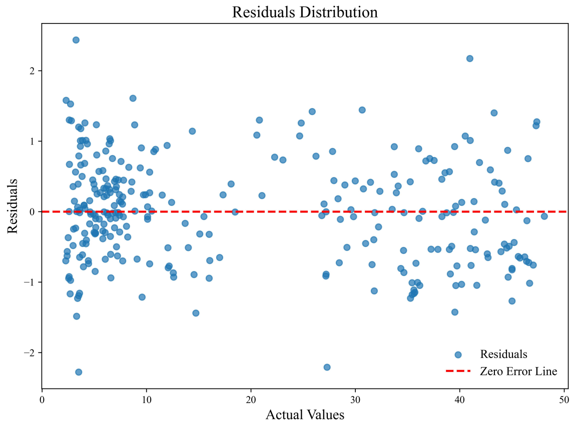
Supplementary Figure 3. Performance comparison of eight machine learning models for Hcj prediction (actual vs. predicted values).



A



B



C

Supplementary Figure 4. Distribution of MLP prediction residuals for in Figure 4A-C. (A) Br; (B) Hcj; (C) (BH)max.