Supplemental Material: Accelerated discovery of cost-effective Nd-Fe-B magnets through adaptive learning

Jie Chen [a], Jian Liu [a], Minjuan Zhang [a, *], Zhanji Dong [b], Zhongjie Peng [b], Xinyi Ji [c], Mei Liu [c], Lanting Zhang [c], Anqi Zhang [d], Hong Zhu [d]

[a] Advanced Engineering Research Division, Toyota Motor Technical Research And Service (Shanghai) Co.,Ltd

[b] Yantai Dongxing Magnetic Materials Inc, 888 Yongda Street, Yantai 265500, China

[c] School of Materials Science and Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

[d] University of Michigan–Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

Supplementary Note 1

Firstly, a correlation analysis is conducted, and 24 features out of 48 are eliminated due to their high correlation coefficient (Pearson coefficient >0.95, see Supplementary Material Fig. S2 for details). Next, a recursive feature elimination (RFE) is performed on the remaining 24 features. The initial data is split into training set and test set, where instead of randomly selecting four samples, we chose all the four ternary systems as test set, which owing to their ternary nature are an indication of the surrogate model's capability to predict the B_r&H_{cj} of multi-doping systems. We adopt average leave-one-out cross-validation root-mean-squared-error on the training set (averaged out over B_r@20℃ and B_r@150℃, hereafter referred to as LOOCV RMSE error, or LOOCV error) as elimination criteria, against which features are retained/eliminated in a greedy fashion. In the meantime, average RMSE error on the test set (hereafter referred to as test error) is monitored and serves as an 'early stopping' mechanism to prevent overfitting to feature subset in RFE process. Starting from 24 features, all possible eliminations are considered, and the "worst" feature, namely, the feature that, upon elimination, gives rise to the best model with the lowest LOOCV error, is eliminated. This is done recursively till there are only two features left.

The RFE process is carried out individually for each ML algorithm, as illustrated using SVR as an example below. Fig. S3 shows the LOOCV error for both B_r@20℃ and B_r@150℃, where the results obtained with Approach 1 is also shown (Compo.) for comparison. LOOCV error of the best model, with the "worst" feature eliminated, as highlighted with blue downward (B_r@20℃) and upward (B_r@150℃) triangles, decreases first when the # features is reduced, and then increases upon further reduction in the # features. The initial decrease is attributed to the elimination of irrelevant/redundant features that contribute little or none to the model accuracy, and rather serve as "noises" that compromise it, the elimination of which hence improves the model, and further elimination of the remaining features that do contribute to the model leads to increased error. The evolution of LOOCV error and test error, as the features are recursively eliminated as described above, is shown in Fig. S3b. LOOCV error reaches a minimum as the # features is reduced to 4, however, test error rises appreciably as the # features is reduced to below 8. Hence, we consider the optimal # features to be 6, where a balance of LOOCV error and test error is achieved. The LOOCV error of the 6-feature set generated with RFE, is significantly lower than that using compositions (Compo.), suggesting that domain-knowledge-based approach is

more advantageous than composition-based approach here. The selected features are shown in Table. S4. Based on the filtered features obtained with RFE, the LOOCV error on the training set for B_r@20℃ and B_r@150℃ are plotted in Fig. S3c for the afore-mentioned six machine learning techniques. SVR achieves the lowest LOOCV error for both B_r@20℃ and B_r@150℃, and is therefore chosen as the machine learning framework for building our surrogate model.

Supplementary Note 2

EI is defined as

$$EI(\mu, \sigma) = \begin{cases} \sigma[\phi(Z) + \Phi(Z)Z], & \sigma > 0 \\ 0, & \sigma = 0 \end{cases}$$

Where Z is defined as

$$Z(\mu, \sigma) = \begin{cases} (\mu - \mu^*)/\sigma, & \sigma > 0 \\ 0, & \sigma = 0 \end{cases}$$

Here μ^* is the maximum value observed so far in the training set, $\phi(Z)$ and $\Phi(Z)$ are the standard normal distribution functions and normal cumulative distribution functions, respectively. As such, EI balances the 'exploration' aspect of a search strategy: $\sigma\phi(Z)$, and 'exploitation' aspect: $\sigma\Phi(Z)Z$. The evaluation of Z (Z-score of μ) requires a reference value, μ^*, taken as the maximum value observed so far in the training set. Note that EI values are generally low when Z < 0 ($\mu < \mu^*$) (unless σ is notoriously large, which is not the case here). Take the optimization of Br@150℃ as an example, μ^* turns out to be very high (10.93 KGs), and as such, due to the cost-performance relationship of Nd-Fe-B (in general, lower cost compositions demonstrate lower performance) investigated in this work, under a cost constraint of, for example, C2 (~10% reduced cost) and C3 (~20% reduced cost), the trained surrogate model would predict values (μ) of all allowed compositions in our target space (Table 1) to be lower than μ^* (10.93 KGs), and hence, very low EI. For example, the predicted range ($\mu \pm \sigma$) and EI of the three compositions in iteration 1 are tabulated below. The E values of the composition 1: C2 and 1:C3 are nearly zero, as the predicted μ is significantly lower than μ^* (10.93 KGs), and this is actually the case for all the allowed compositions in the target space. The near-zero values of EI defeat the purpose of using EI. Although this could be remedied by defining different μ^* for different cost constraint, the choice of μ^* is arbitrary. We therefore turn to a simpler acquisition function, UCB, defined as $\mu + k\sigma$, with k=1. Here k is a pre-determined term that determines the trade-off between 'exploration' and 'exploitation' aspect of the search strategy. As such, although differing in their functional form, both EI and UCB trades-off 'exploration' and 'exploitation', and could be used to guide experimental design. Note that UCB has also proven successful in Bayesian search (Attia. et al., Nature 578, 397–402 (2020) [1]). We also note that for many practical problems, the choice of the surrogate model matters more than the acquisition function, and actually the performance of EI and UCB has been shown to be close in terms of the rate at which 'high-performance' materials are successfully 'mined' (Liang et al., npj Computational Materials (2021) 7:188 [2]).

Iteration	Predicted Br@150℃ (KGs)	EI (KGs)
1:C1	11.24±0.37	0.35
1:C2	9.81±0.77	0.03
1:C3	9.62±0.76	0.01

Supplementary Note 3

We added a wiggle room to the cost constraints to promote exploration. This wiggle room was set to be 2% for Stage 1 and 1% for Stage 2. For example, a ~20% cost reduction level with a wiggle room of 2% means that the cost reduction is required to be at least 18%.

Supplementary Note 4

The temperature coefficient of B_r (α_{Br}) and H_{cj} (β_{Hcj}) is calculated as the percentage change in B_r (or H_{cj}) per °C, as compared to the value at 20°C, as shown below.

$$\alpha_{Br}|_T = \frac{Br|_T - Br|_{20°C}}{T - 20} * 100$$

$$\beta_{Hcj}|_T = \frac{Hcj|_T - Hcj|_{20°C}}{T - 20} * 100$$

Where the relative change is multiplied by 100 and hence converted to percentage reduction. As is shown in Fig. S7, B_r and H_{cj} do not usually vary linearly with T, and therefore α_{Br} and β_{Hcj} are temperature-dependent, and are usually specified at a given temperature, hence $\alpha_{Br}|_T$ and $\beta_{Hcj}|_T$. In our case, both α_{Br} and β_{Hcj} are calculated at 150°C. For example, α_{Br} of pristine Nd-Fe-B at 150°C is -0.133, meaning that B_r shows an average reduction of 0.133%/°C as temperature is increased from 20°C to 150°C.

Table. S1 Composition (main phase wt%) and magnetic properties (B_r, H_{cj}) of the initial dataset comprising of 24 samples.

Sample	La	Ce	Y	Co	Ni	B_r: 20	B_r: 80	B_r: 150	H_{cj}: 20	H_{cj}: 150	Cost	Relative cost
1	0	0	0	0	0	13.24	12.38	10.93	15.48	3.61	236.331	0
2	0	0	0	18.61	0	12.81	12.1	11.33	14.68	2.44	306.891	0.299
3	0	5.77	0	0	0	12.5	11.36	9.862	11.78	2.59	182.181	-0.229
4	0	5.32	0	18.05	0	11.98	11.42	10.6	12.98	1.82	252.741	0.069
5	0	11.86	0	0	0	11.34	10.41	8.537	5.97	0.441	128.031	-0.458
6	4.37	0	0	17.93	0	12.39	11.77	10.99	7.705	2.12	253.266	0.072
7	0	0	0	36.61	0	11.52	10.16	7.73	0.243	0.084	377.451	0.597
8	0	11.2	0	18.74	0	10.9	10.25	9.31	4.96	0.502	198.591	-0.16
9	4.5	0	0	0	0	12.09	11	9.73	12.72	2.85	182.706	-0.227
10	4.47	0	0	36.99	0	11.03	10.59	9.98	11.3	3.09	323.826	0.37
11	7.93	0	0	16.54	0	12.01	11.44	10.67	3.56	1.27	199.641	-0.155
12	9.2	0	0	30.54	0	11.34	10.83	10.2	3.86	0.932	270.201	0.143
13	0	0	0	0	1.75	12.21	11.4	10.18	17.23	3.86	240.023	0.016
14	0	5.71	0	0	1.91	11.83	10.96	10.33	10.05	2.81	185.873	-0.214
15	0	0	4.55	0	0	12.21	11.35	10.1	13.35	3.32	210.331	-0.11
16	3.61	0	2.66	0	0	9.76	8.82	7.29	3.5	0.742	169.706	-0.282
17	0	5.08	2.38	0	0	12.1	11.05	9.58	11.35	2.35	169.181	-0.284
18	0	0	4.9	9.11	0	11.75	11.11	10.22	14.57	3.51	245.611	0.039
19	0	0	4.56	0	1.6	10.66	9.91	8.85	12.72	2.8	214.023	-0.094
20	4.37	6.13	0	0	0	10.37	9.525	7.081	3.093	0.162	128.556	-0.456
21	3.7	0	2.33	8.66	0	10.63	10.01	9.21	5.062	1.96	204.986	-0.133
22	0	4.99	2.38	9.15	0	11.81	11.11	10.11	11.62	2.13	204.461	-0.135
23	4.4	5.85	0	9.26	0	11.24	10.54	9.63	6.62	1.92	163.836	-0.307
24	4.63	6	0	0	1.81	10.78	9.7	8.3	3.589	0.913	132.248	-0.44

Table. S2 Prices of the doping species in unit of RMB/KG. The values were retrieved from [3] on 2021/9/1.

Element	Nd/Pr	La	Ce	Y	Fe	Co	Ni
Price (RMB/KG)	750	35	28	230	8	400	150

Table. S3 Atomic features for domain-knowledge-based fingerprinting.

Atomic feature	La	Ce	Y	Co	Ni
Atomic number Z_R (Z_T)	57	58	39	27	28
Atomic radii r_R (r_T)	180	163	163	111	110
Electronegativity χ_R (χ_T)	1.10	1.12	1.22	1.88	1.91
Atomic mass m_R (m_T)	138.91	140.12	88.91	58.93	58.69
Valence electron number VEN_R (VEN_T)	3	4	3	9	10
Spin angular moment S_{4f} (S_{3d})	0.0	0.5	0.0	1.5	1.0
Orbital angular moment L_{4f} (L_{3d})	0.0	3.0	0.0	3.0	3.0
Ionization potential IP_R (IP_T)	538.09	534.39	599.87	760.4	737.14
Density ρ_R (ρ_T)	6.146	6.689	4.472	8.9	8.908
Melting point Tm_R (Tm_T)	1193	1068	1799	1768	1728
Boiling point Tb_R (Tb_T)	3743	3633	3609	3200	3186
Enthalpy of fusion Ef_R (Ef_T)	6.2	5.5	11.4	16.2	17.2
Enthalpy of vaporization Ev_R (Ev_T)	400	350	380	375	378
Enthalpy of atomization Ea_R (Ea_T)	431	423	425	426	431
Curie Temperature Tc_R (Tc_T)	**530**	424	565	1388	628
Saturation magnetization μ_R (μ_T)	**13.8**	11.7	14.1	1.72	0.60

Table. S4 Selected feature with the six different ML algorithms, and the corresponding LOOCV error (KGs) in this work.

Estimator	Selected Features	B_r@20℃: CV	B_r@150℃: CV
GP	['IP_{RT}', 'x_{RT}', 'Ev_R', 'Ef_{RT}']	0.6	0.83
SVR	['Tc_T', 'Tc_{RT}', 'IP_{RT}', 'μ_{RT}', 'Ev_R', 'IP_T']	0.39	0.61
Ridge	['Tc_{RT}', 'IP_T', 'Ev_R', 'Ea_{RT}']	0.50	0.89
KNN	['$L4f_R$', 'Ea_{RT}', 'IP_T', 'm_R', 'Ef_R']	0.62	0.83
ANN	['μ_T', 'μ_R', 'μ_{RT}', 'r_R', 'm_R', 'IP_T']	0.57	1.07
DT	['$L4f_R$', 'Tm_{RT}', 'Tc_{RT}']	0.74	0.85

Table. S5 Recommended compositions and the corresponding performance during Stage 1, aimed at maximizing B_r@150℃. Performance of the pristine Nd-Fe-B (Iteration 0) is also included as benchmark.

	Composition	**B_r@20℃ (KGs)**		**B_r@150℃ (KGs)**		
Iteration	**Main phase wt%**	**Predicted**	**Actual**	**Predicted**	**Actual**	**Relative Cost**
0	Pristine Nd-Fe-B	---	13.24	---	10.93	---

1:C1	Ce_3Co_7	12.85±0.30	12.66	11.24±0.37	10.78	0%
1:C2	$La_6Co_{10}Ni_3$	11.45±0.64	11.48	9.81±0.77	10.15	-11.7%
1:C3	$La_6Ce_2Co_{10}Ni_3$	11.28±0.64	10.43	9.62±0.76	8.99	-19.6%
2:C1	$La_3Co_9Ni_{0.5}$	12.39±0.22	12.45	11.03±0.28	10.78	0%
2:C2	$La_5Co_{10}Ni_{1.5}$	11.96±0.25	12.15	10.72±0.33	10.55	-8.4%
2:C3	$La_7Co_{10}Ni_2$	11.78±0.29	12.00	10.36±0.37	10.51	-18.5%

Table. S6 Recommended compositions and the corresponding performance during Stage 2, aimed at maximizing B_r@80°C*H_{cj}@150°C. Performance of the pristine Nd-Fe-B (Iteration 0) is also included as a baseline for comparison.

	Composition	**B_r@80°C (KGs)**		**H_{cj}@150°C (KOe)**		**$M_{Br\&Hcj}$ (KGs*KOe)**		
Iteration	**Main phase wt%**	**Predicted**	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**	**Relative Cost**
0	Pristine Nd-Fe-B	---	12.38	---	3.61	---	44.69	---
3:C1	Ce_1Co_1	12.04±0.29	11.99	3.76±0.37	4.31	45.33±5.66	51.68	-2.4%
3:C2	$La_1Ce_1Ni_{0.25}$	11.85±0.24	11.80	3.73±0.35	3.89	44.18±4.93	45.90	-9.2%
3:C3	$Ce_5Y_{0.5}Co_1$	11.45±0.15	11.34	3.15±0.30	3.27	36.04±3.68	37.08	-19.7%

Table. S7 Composition (nominal wt% and main phase) and the wt% of Nd/Pr replaced for all the compositions recommended in the iterative design process.

Iteration	**Main phase wt%**	**Nominal wt%**	**% Nd/Pr**
1:C1	Ce_3Co_7	$Ce_{4.0}Co_{7.2}$	13.0%
1:C2	$La_6Co_{10}Ni_3$	$La_{10.5}Co_{10.2}Ni_{4.4}$	34.1%
1:C3	$La_6Ce_2Co_{10}Ni_3$	$La_{10.5}Ce_{2.6}Co_{10.2}Ni_{4.4}$	42.5%
2:C1	$La_3Co_9Ni_{0.5}$	$La_{5.3}Co_{9.2}Ni_{0.7}$	17.2%
2:C2	$La_5Co_{10}Ni_{1.5}$	$La_{8.8}Co_{10.2}Ni_{2.2}$	28.6%
2:C3	$La_7Co_{10}Ni_2$	$La_{12.3}Co_{10.2}Ni_{3.0}$	39.9%
3:C1	Ce_1Co_1	$Ce_{1.33}Co_{1.03}$	4.3%
3:C2	$La_1Ce_1Ni_{0.25}$	$La_{1.76}Ce_{1.33}Ni_{0.37}$	10.0%
3:C3	$Ce_5Y_{0.5}Co_1$	$Ce_{6.62}Y_{0.53}Co_{1.03}$	23.2%

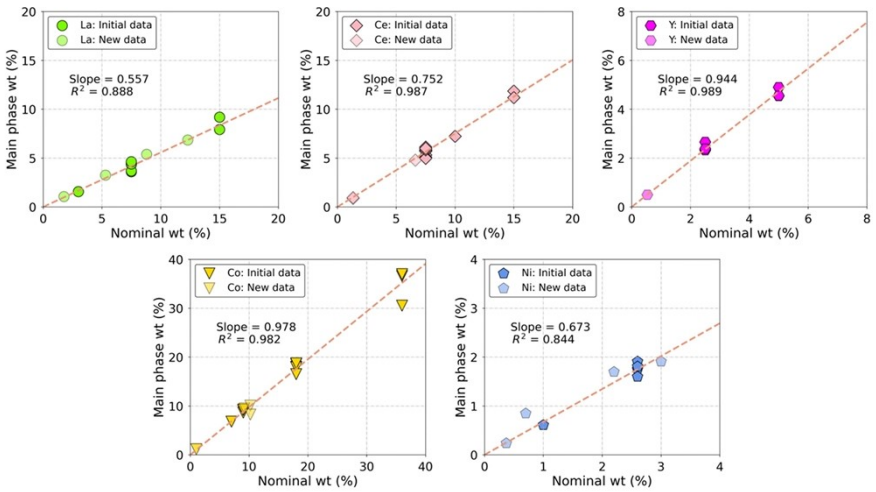

Fig. S1 Variation of main phase wt% vs. nominal wt% for the five doping species, based on the augmented dataset (24 samples in the initial data, plus 9 new compositions). A linear fit renders a slope (intercept is forced to be 0), representing the average ratio of main phase wt% and nominal wt%.

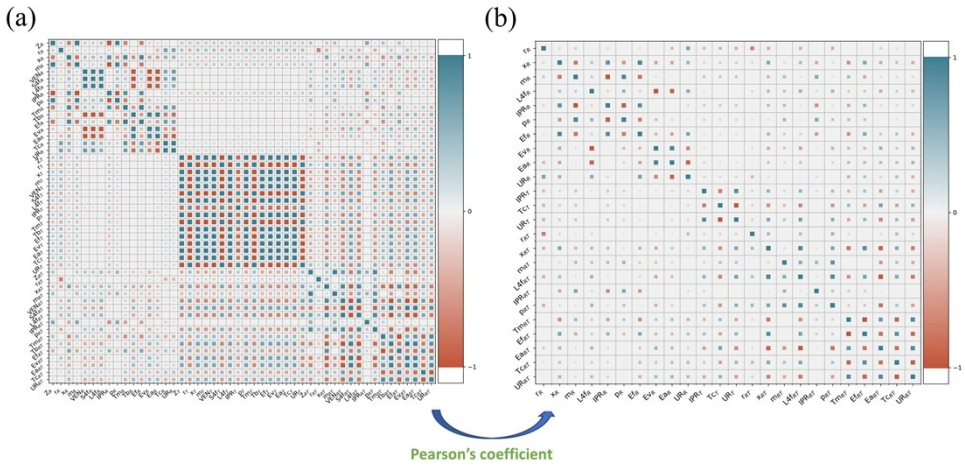

Fig. S2 Heatmap showing the Pearson's correlation coefficients among the atomic features: (a) before and (b) after the filtering process. The number of features is reduced from 48 to 24, with a cutoff of 0.95 for feature elimination.

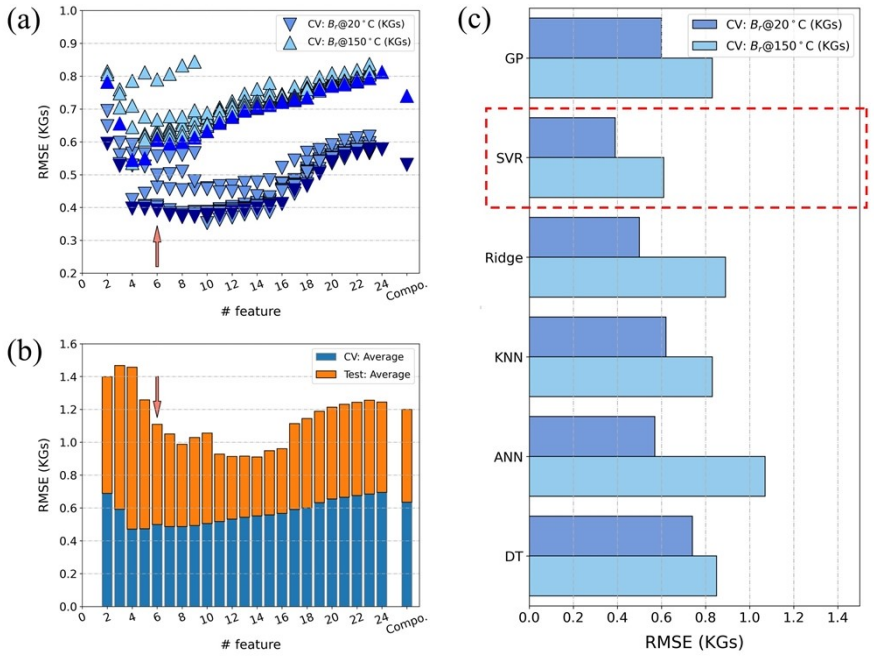

Fig. S3 (a) Variation of LOOCV error and test error with # features during the RFE process for SVR, (b) Bar-plot showing LOOCV error and test error for composition-based approach (Compo.) and domain-knowledge-based approach during the RFE process for SVR, (c) LOOCV error on the train/test set for B_r@20°C and B_r@150°C, evaluated for six different machine learning techniques. The salmon-colored arrow in (a) and (b) indicates the # features chosen (# features = 6) for SVR.

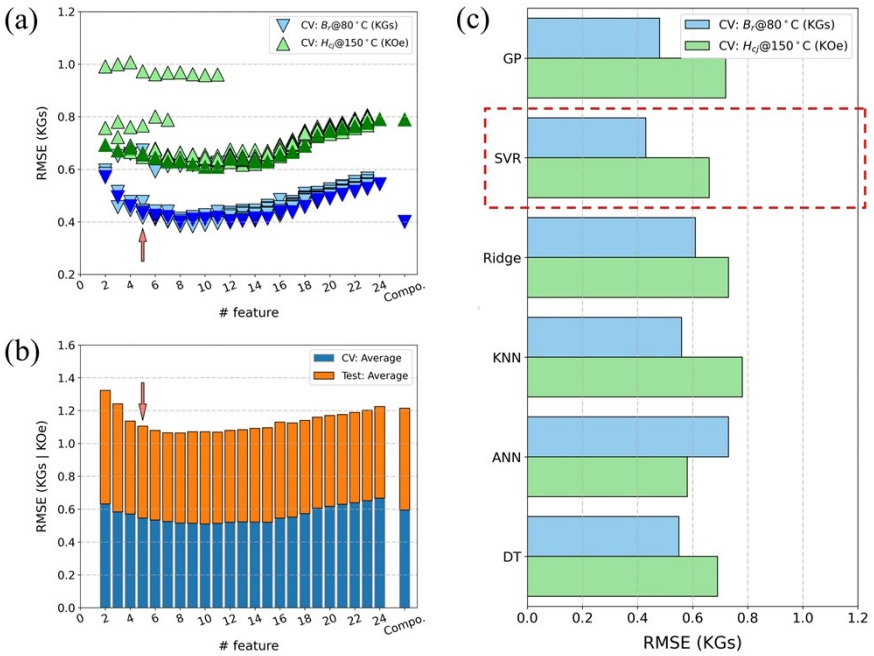

Fig. S4 (a) Variation of LOOCV error and test error with # features during the RFE process for SVR, (b) Bar-plot showing LOOCV error and test error for composition-based approach (Compo.) and domain-knowledge-based approach during the RFE process for SVR, (c) LOOCV error on the train/test set for B_r@80°C and H_{cj}@150°C, evaluated for six different machine learning techniques. The salmon-colored arrow in (a) and (b) indicates the #

features chosen (# features = 5) for SVR. The results are shown for Stage 2.

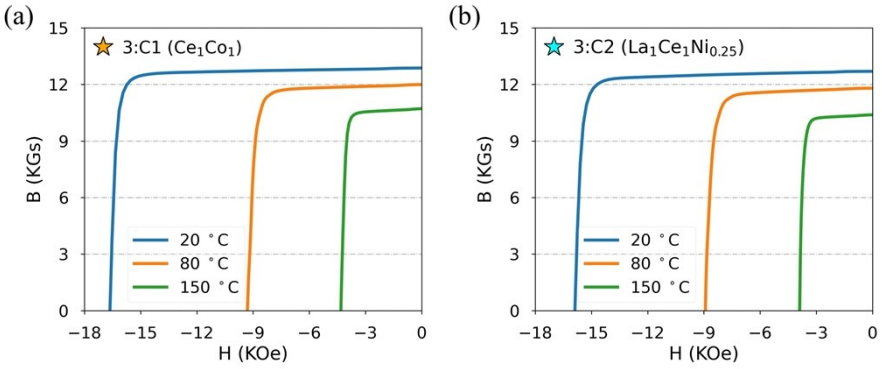

Fig. S5 Demagnetizing curves (20℃, 80℃, and 150℃) of the first two compositions recommended in iteration 3: (a) **3: C1** (Ce_1Co_1), (b) **3: C2** ($La_1Ce_1Ni_{0.25}$).

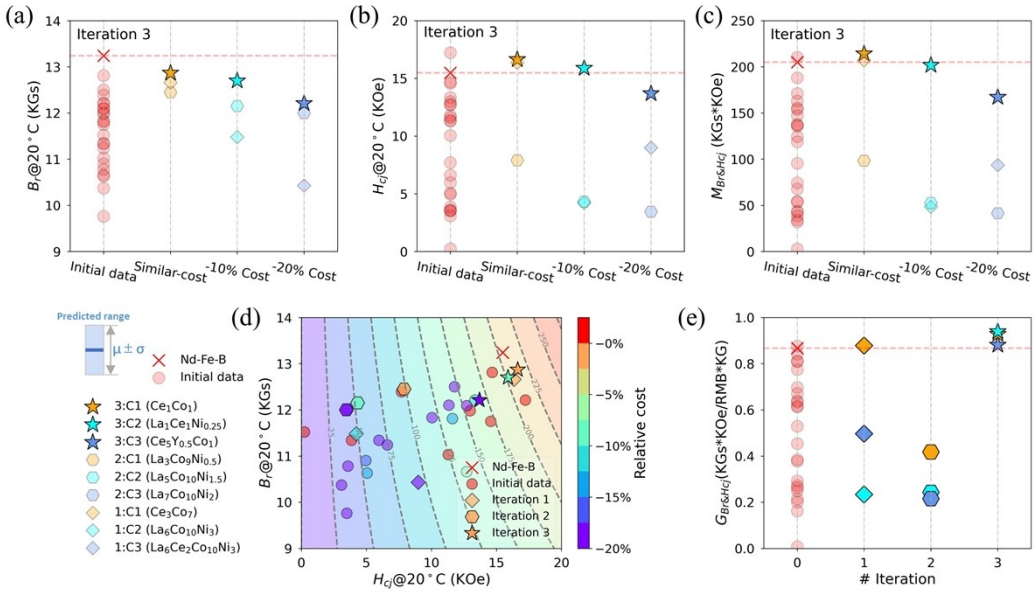

Fig. S6 Recommended compositions and their corresponding (a) B_r@20℃, (b) H_{cj}@20℃, (c) $M_{Br\&Hcj}$ in iteration 3. (d) Contour plot showing the performance of recommended compositions as compared to initial dataset (with pristine Nd-Fe-B highlighted as cross), with symbols colored based on the relative cost of the corresponding compositions w.r.t. pristine Nd-Fe-B. The dashed gray contour lines represent the variation of $M_{Br\&Hcj}$. (e) Variation of $G_{Br\&Hcj}$ over iterations, with the initial dataset shown as iteration 0. In (a)-(c) model-predicted range (μ±σ) is shown as vertical bars for iteration 3, similar to that in Fig. 5. Note that $M_{Br\&Hcj}$ is defined as B_r@20℃*H_{cj}@20℃ here to reflect the overall performance at RT, and $G_{Br\&Hcj}$ defined as $M_{Br\&Hcj}$/cost.

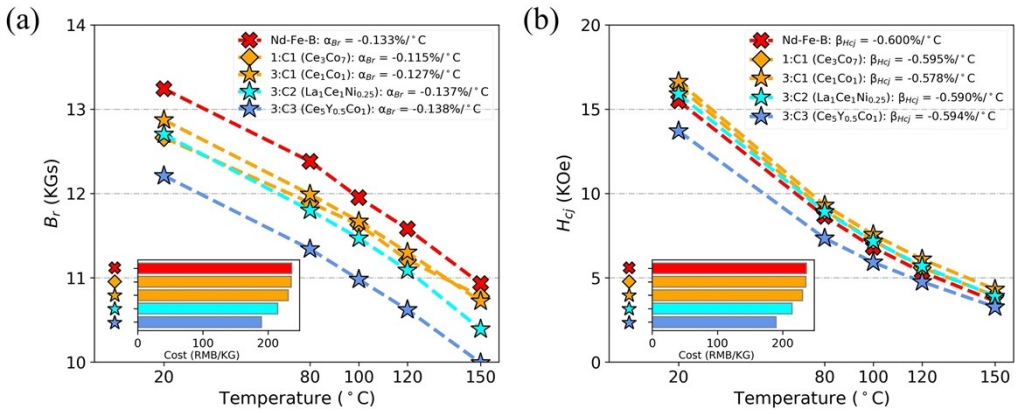

Fig. S7 Temperature dependence of (a) B_r, (b) H_{cj} for selected compositions with positive performance-cost ratio gain $\Delta G_{Br\&Hcj}$, along with that of pristine Nd-Fe-B. Costs of the selected compositions are also shown in the inset bar plot.

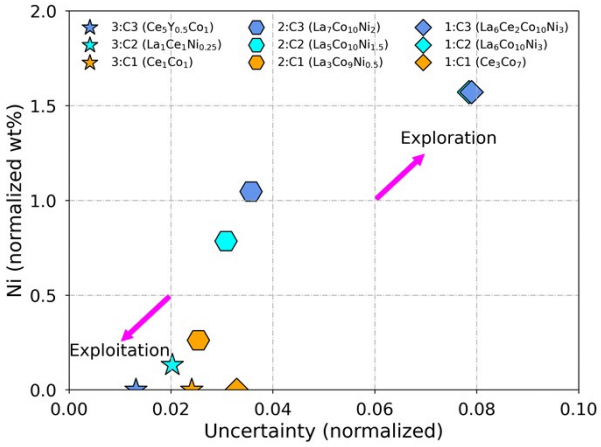

Fig. S8 Variation of Ni wt% and uncertainty for the recommended compositions throughout the iterative design process. Ni wt% is normalized w.r.t. that of the highest Ni wt% in the initial dataset, and hence a value greater than 1 indicates extrapolation in terms of Ni wt% during the iterative design. Uncertainty is calculated as predicted uncertainty/predicated mean, and calculated based on B_r@150°C for iteration 1-2 and B_r@80°C for iteration 3. In this context, a high Ni wt% along with high uncertainty indicate an 'explorative' nature (iteration 1), and the opposite indicates an 'exploitative' nature (iteration 2-3).

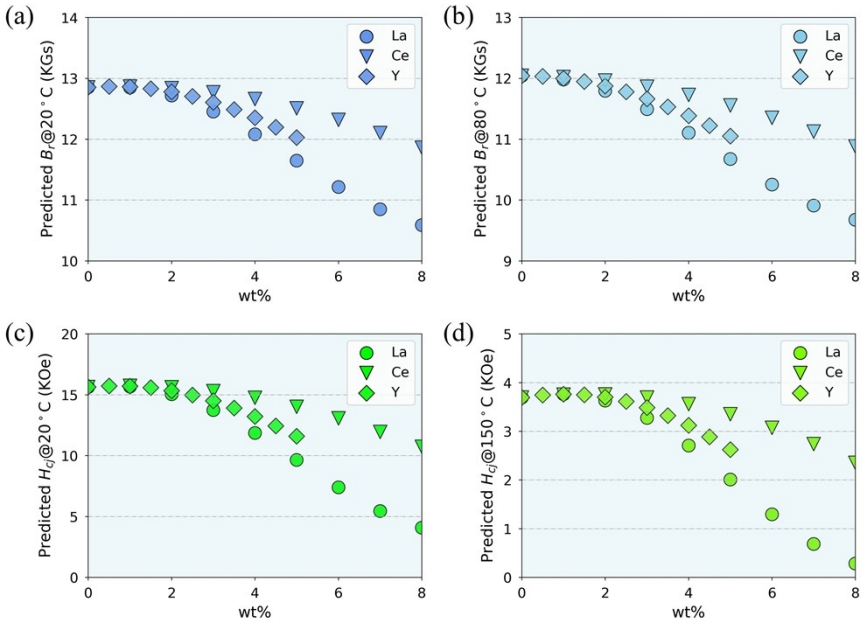

Fig. S9 Variation of (a) B_r@20℃, (b) B_r@80℃, (c) H_{cj}@20℃, (d) H_{cj}@150℃ with wt% of La/Ce/Y as predicted by the model (Stage 2).

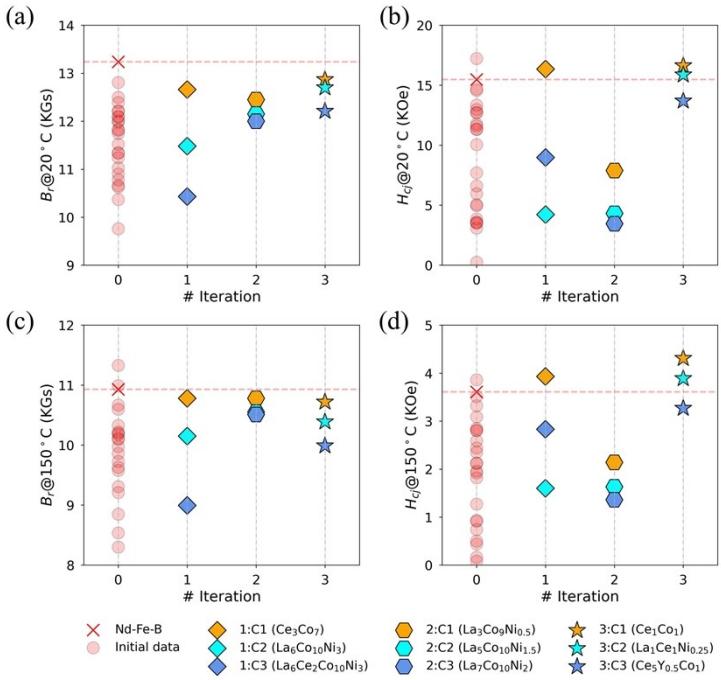

Fig. S10 Variation of (a) B_r@20℃, (b) H_{cj}@150℃, (c) B_r@150℃, (d) H_{cj}@150℃ with the # iteration during the adaptive learning process.

Reference

1. P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B.

Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon and W. C. Chueh, *Nature*, 2020, **578**, 397-402.

2. Q. Liang, A. E. Gongora, Z. Ren, A. Tiihonen, Z. Liu, S. Sun, J. R. Deneault, D. Bash, F. Mekki-Berrada, S. A. Khan, K. Hippalgaonkar, B. Maruyama, K. A. Brown, J. Fisher Iii and T. Buonassisi, *npj Computational Materials*, 2021, **7**.
3. https://price.metal.com/.