Supplementary Materials

S.1 REE datasets curated for this study

Table S1: A representative subset (10%) of the REE dataset adopted in this study**.**

Sample	Input feature [morlar%]										Label [ppm, by molar]																
	SiO_2	Al_2O_3	CaO	Fe_2O_3	MgO	Na_2O	K_2O	TiO_2	P_2O_5	MnO	CeO_2	Dy_2O_3	Er_2O_3	EuO	Gd_2O_3	Ho_2O_3	La_2O_3	Lu_2O_3	Nd_2O_3	Pr_2O_3	Sc_2O_3	Sm_2O_3	Tb_2O_3	Tm_2O_3	Yb_2O_3	Y_2O_3	Total
1	44.80	14.36	26.88	2.50	7.41	1.53	0.46	1.19	0.44	0.03	58.62	2.02	1.10	1.21	2.29	0.37	16.44	0.14	12.60	3.44	21.52	2.31	0.33	0.14	0.94	18.64	142.1
2	34.26	11.54	27.43	2.41	8.74	13.29	0.34	1.01	0.57	0.02	46.67	1.58	0.89	1.00	1.78	0.29	12.61	0.11	9.97	2.70	18.88	1.86	0.27	0.11	0.74	14.41	113.9
3	67.51	14.34	5.76	6.57	1.69	0.97	1.84	1.00	0.09	0.05	58.54	2.34	1.33	1.09	2.45	0.43	15.64	0.18	12.63	3.39	24.51	2.53	0.37	0.18	1.14	21.53	148.3
4	72.21	16.02	1.80	5.27	1.72	0.32	1.09	1.26	0.09	0.02	81.41	2.87	1.71	1.42	3.32	0.58	20.00	0.21	17.37	4.60	37.51	3.47	0.49	0.23	1.46	29.41	206.1
5	68.79	19.30	1.41	4.37	2.73	0.37	1.38	1.17	0.17	0.03	98.13	3.61	2.19	1.70	4.21	0.71	24.65	0.28	22.08	5.83	42.63	4.48	0.66	0.30	1.84	37.21	250.5
6	66.78	19.58	2.06	4.97	2.82	0.37	1.51	1.30	0.28	0.04	108.75	3.73	2.12	1.93	4.42	0.73	26.74	0.27	24.29	6.39	39.06	4.74	0.67	0.29	1.84	37.22	263.2
7	79.28	9.55	2.16	4.19	2.12	1.22	0.62	0.54	0.08	0.04	25.71	0.91	0.53	0.54	1.14	0.18	6.27	0.08	5.83	1.55	25.57	1.20	0.16	0.08	0.48	9.44	79.7
8	38.33	9.84	35.57	2.04	11.27	1.27	0.27	1.02	0.00	0.00	37.94	1.25	0.67	0.93	1.46	0.23	10.09	0.08	8.08	2.13	12.81	1.57	0.20	0.09	0.57	11.58	89.7
9	33.32	11.56	34.57	7.94	10.05	0.20	0.27	0.76	0.05	0.82	55.45	2.95	1.87	1.48	3.21	0.60	12.75	0.22	14.06	3.56	28.82	3.07	0.50	0.24	1.53	37.14	167.5
10	43.20	10.86	29.38	2.42	9.51	1.17	0.09	2.65	0.12	0.05	45.03	1.77	1.09	0.95	1.94	0.35	12.05	0.15	9.45	2.55	31.98	1.92	0.31	0.15	1.01	17.73	128.4

Table S2: **A separate REE dataset consisting of 12 coal fly ashes from Poland.** This dataset is curated based on the raw data reported by Franus et al. (*88*).

Sample	Input feature [morlar%]										Label [ppm, by molar]															
	SiO_2	Al_2O_3	CaO	Fe_2O_3	MgO	Na_2O	K_2O	TiO_2	P_2O_5	MnO	CeO_2	Dy_2O_3	Er_2O_3	EuO	Gd_2O_3	Ho_2O_3	La_2O_3	Lu_2O_3	Nd_2O_3	Pr_2O_3	Sc_2O_3	Sm_2O_3	Tb_2O_3	Tm_2O_3	Yb_2O_3	Y_2O_3	Total
1	65.44	19.32	3.79	2.77	4.20	0.88	2.21	1.06	0.24	0.07	64.93	1.89	1.00	1.17	2.10	0.40	15.84	0.14	13.69	3.60	24.64	2.72	0.35	0.16	0.97	20.14	153.8
2	65.42	18.93	4.02	2.81	4.30	0.94	2.25	0.97	0.28	0.06	63.97	1.93	1.06	1.20	2.13	0.37	15.81	0.16	13.95	3.64	1.64	2.58	0.37	0.17	1.03	20.35	130.3
3	63.68	18.93	5.41	2.86	4.65	1.01	2.06	1.06	0.24	0.08	68.54	2.04	1.19	1.33	2.32	0.40	16.88	0.15	14.98	3.91	25.95	2.88	0.42	0.17	1.09	22.69	164.9
4	62.44	16.23	6.54	4.00	6.53	0.95	2.08	0.92	0.16	0.14	56.60	1.81	0.97	0.96	1.88	0.35	14.10	0.14	12.29	3.19	19.99	2.46	0.31	0.13	0.87	18.96	135.0
5	64.74	19.15	3.82	2.88	4.46	1.26	2.37	0.99	0.24	0.09	66.69	2.10	1.11	1.28	2.26	0.38	16.65	0.15	13.94	3.72	22.32	2.84	0.37	0.17	0.96	20.82	155.8
6	65.35	23.78	1.56	2.44	2.41	0.60	2.30	1.30	0.23	0.02	92.74	2.82	1.67	1.89	3.51	0.59	22.15	0.22	21.23	5.48	37.70	4.26	0.57	0.24	1.47	31.01	227.6
7	61.36	20.08	5.02	2.52	5.24	1.54	2.63	1.05	0.48	0.06	70.52	2.20	1.21	1.39	2.53	0.43	17.26	0.17	15.26	4.03	26.38	3.10	0.41	0.19	1.09	23.85	170.0
8	62.18	16.44	6.53	3.83	6.56	1.28	1.94	0.88	0.21	0.12	57.33	1.73	0.97	1.05	1.94	0.36	14.11	0.14	12.42	3.18	20.55	2.56	0.33	0.15	0.89	19.02	136.7
9	52.71	16.28	13.29	4.64	8.03	1.79	1.90	0.94	0.29	0.12	58.74	1.95	1.06	1.17	2.05	0.37	14.79	0.15	12.24	3.31	22.97	2.62	0.38	0.15	0.95	21.90	144.8
10	42.18	18.50	29.67	2.01	3.03	1.67	0.83	1.97	0.09	0.03	69.56	1.56	0.74	1.15	1.94	0.29	19.06	0.10	14.03	3.99	15.79	2.60	0.36	0.12	0.67	15.61	147.6
11	47.70	11.70	35.30	1.94	1.95	0.17	0.14	1.01	0.05	0.04	68.22	2.02	1.08	1.19	2.43	0.39	18.49	0.15	16.13	4.22	16.45	2.88	0.39	0.17	0.99	23.93	159.1
12	52.36	2.39	32.67	2.42	9.03	0.15	0.14	0.47	0.08	0.29	13.42	0.49	0.33	0.23	0.56	0.11	3.42	0.05	2.70	0.72	4.77	0.57	0.09	0.05	0.32	6.17	34.0

S.2 Effect of the hidden layer number of the neural network

To showcase the optimization of the number of hidden layers of the neural network, here we compare the model accuracy across three settings while maintaining the number of hidden neurons as a constant for fairness. The first model comprises a single layer with 26 neurons, the second model incorporates two layers with 20 and 6 neurons, respectively, and the third model features three layers with 17, 6, and 3 neurons, respectively. Each model undergoes independent hyperparameter optimization using a grid search within their typical ranges, as outlined in Sec. 2.3.

Upon evaluation, the single-layer model demonstrates inferior accuracy on both the training and test sets. The three-layer model attains the highest accuracy on both sets; however, a noticeable disparity emerges between the training and test accuracy, accompanied by a considerably larger standard deviation in the test R^2. These observations suggest a tendency of overfitting in the three-layer model. In contrast, the two-layer model emerges as the optimal choice, reflected in the lower bound of the shadowed region (i.e., mean – standard deviation). Consequently, for this study, we adopt the neural network configuration featuring two hidden layers.

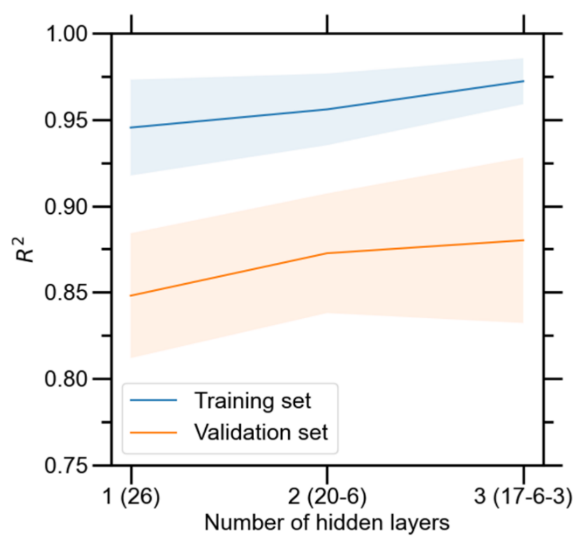
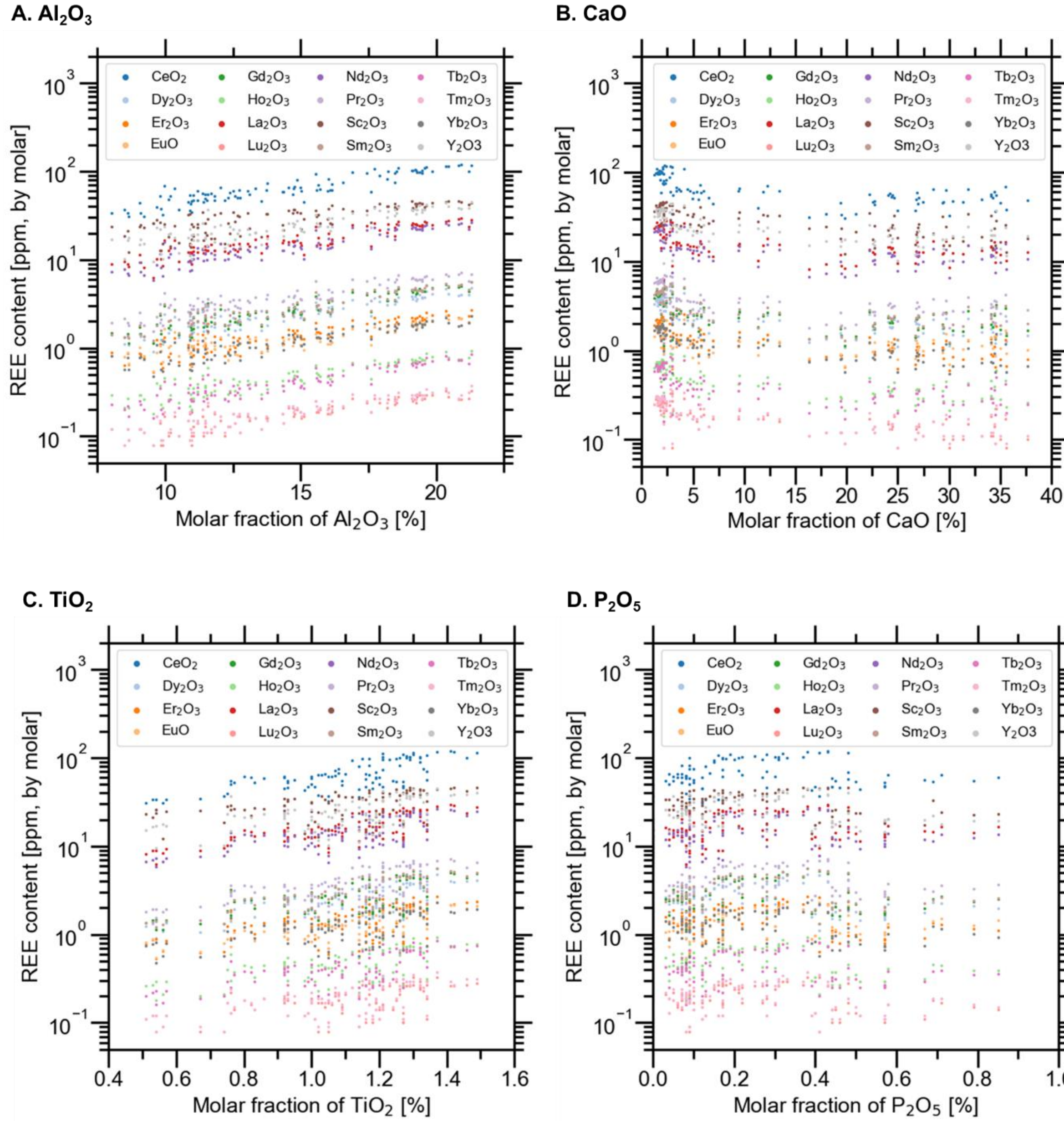


Figure S1: Prediction accuracy (as captured by the coefficient of determination, R^2) offered by three neural network models of different numbers of hidden layers. For illustrative purposes, the total number of neurons is kept the same as 26 across the three compared cases. The results are presented for both the training (80%) and validation (20%) sets, obtained through stratified sampling. The shaded areas represent the standard deviation of the prediction accuracy on the two sets, calculated from 30 independent repeats of the model training.

S.3 Correlation between the contents of major oxides and REEs



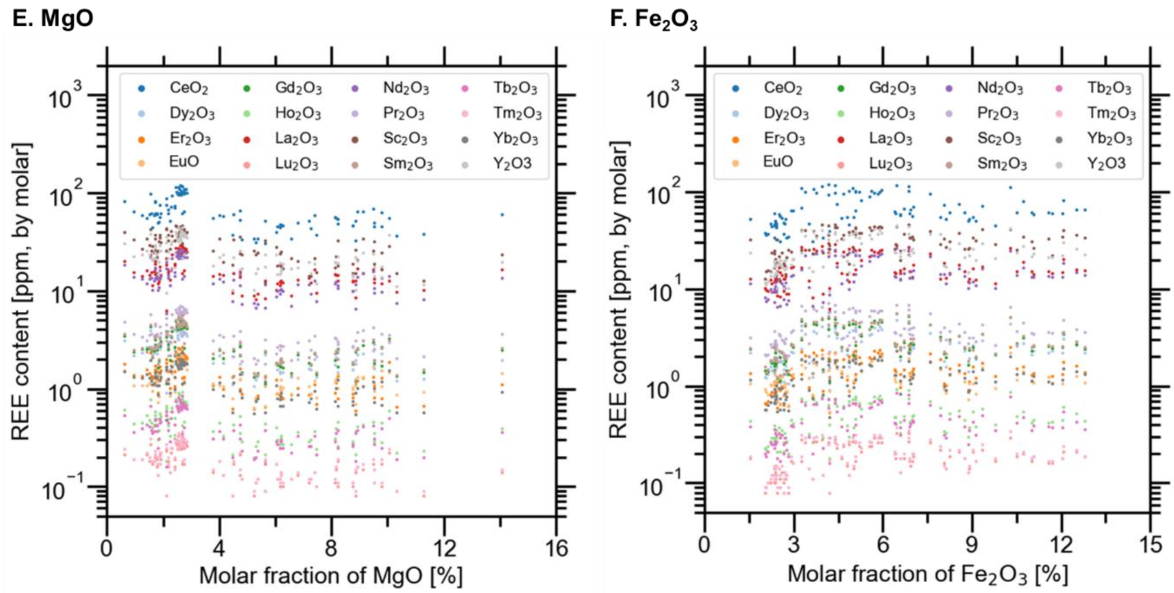


Figure S2: **Correlation between the contents of the selected oxides (molar fraction, as the model input) and individual REEs (ppm by molar, as the model output)**, regarding **(A)** Al_2O_3**, (B)** CaO, **(C)** TiO_2, **(D)** P_2O_5, **(E)** MgO, and **(F)** Fe_2O_3.

S.4 Comparision of the prediction accuracy between different single-task models

Table S3: **Comparison of the accuracy of the different single-task models in predicting the total REE content.** The training (80%) and test (20%) sets are obtained with stratified sampling. The results are reported based on the coefficient of determination, R^2, and mean absolute percentage error, MAPE; and the mean and standard deviation are obtained from 30 repetitions. For each model, the hyperparameters are optimized by using a grid search for the optimal R^2 within their typical ranges. Unless specified, the hyperparameters of the individual models are kept as the default values in their respective libraries.

ML model		Linear regression (40)	Random forest (40)	SVM (40)	XGboost (65)	Neural network (64)
Key hyperparameters		N/A	estimators= 30 max_depth = 4	degree=2 C=0.15	n_estimators=30 max_depth=4 eta=0.2	hidden layers = (20, 6) epoch = 500 alpha = 0.001 learning rate$10^{-3.33}$ weight decay = 10^{-2}
Training set	R^2	0.876 ± 0.010	0.946 ± 0.005	0.896 ± 0.007	0.998 ± 0.001	0.940 ± 0.019
	MAPE [%]	9.8 ± 0.5	13.9 ± 0.6	19.2 ± 0.6	2.3 ± 0.3	6.7 ± 1.1
Test set	R^2	0.847 ± 0.060	0.859 ± 0.057	0.866 ± 0.024	0.877 ± 0.056	0.875 ± 0.037
	MAPE [%]	10.2 ± 2.0	22.1 ± 4.4	22.0 ± 1.9	20.6 ± 4.3	9.5 ± 1.3

S.5 Core code of the multi-task neural network model

The core code and trained multi-task neural network models are available in the below link: https://github.com/Little-uni/multi-task-neural-network