Accelerated Prediction of Perovskite Material Properties with Classical Machine Learning and Graph Neural Network

Zhihao Dong, Yujin Ji, Youyong Li*

Institute of Functional Nano & Soft Materials (FUNSOM), Jiangsu Key Laboratory for Carbon-Based Functional Materials & Devices, Soochow University, Suzhou, Jiangsu 215123, PR China

*Corresponding author: Youyong Li, E-mail: ligroup2023@vip.163.com

Supporting Information

Table S1. Feature ID and Attribute Names for Pearson Correlation Analysis.

No	Property	Unit
1	Atomic number	
2	Atomic radius	pm
3	Atomic radius by Rahm	pm
4	Atomic volume	cm^3/mol
5	Atomic weight	
6	Dipole polarizability	a.u.
7	Electron affinity	eV
8	Number of electrons	
9	Pauling's scale of electronegativity	
10	Group in periodic table	
11	First ionization energy	eV
12	Mass number	
13	Mendeleev's number	
14	Metallic radius with 12 nearest neighbors	pm
15	Molar heat capacity @ 25 C, 1 bar	J/(mol K)
16	Number of neutrons	Number of neutrons
17	Period in periodic table	Period in periodic table
18	Pettifor scale	
19	Proton affinity	kJ/mol
20	Number of protons	
21	Van der Waals radius	

Table S1. Cross-validation MAE errors of formation energy

No	ANN	SVR	RF	GBDT
1	0.367114	0.271016	0.251597	0.242682
2	0.375722	0.26937	0.252829	0.245718
3	0.36726	0.274309	0.255171	0.243964
4	0.384651	0.278298	0.257878	0.249249
5	0.363926	0.275315	0.254205	0.246797

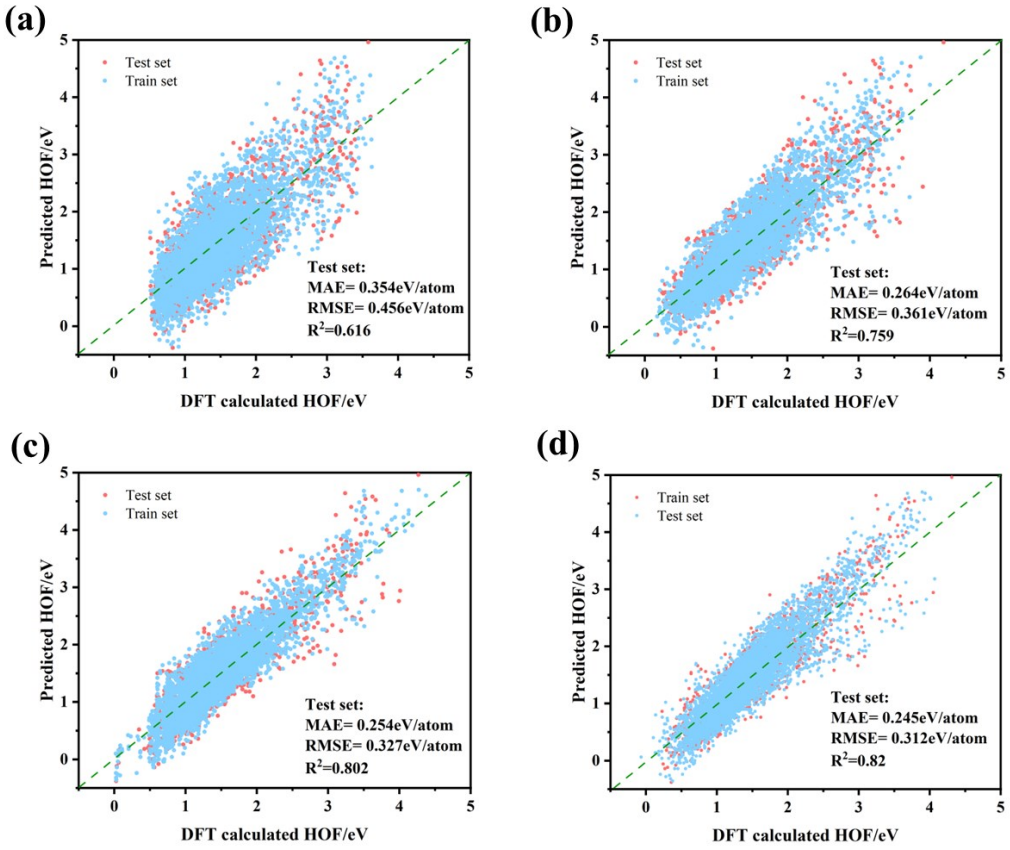

Figure S1. Scatterplot of predictions of perovskite formation energy by classical machine learning models. (a). ANN (b). SVR (c). RF (d) GBDT

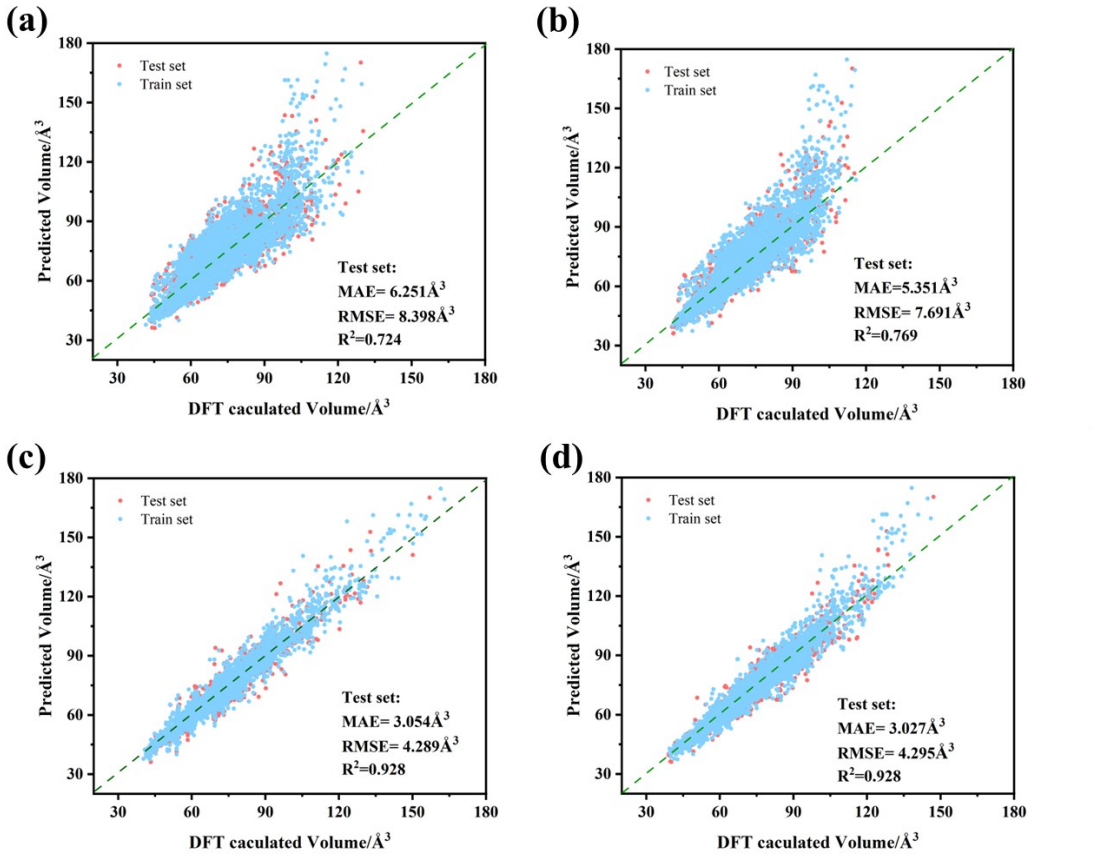

Figure S2. Scatterplot of predictions of perovskite volume by classical machine learning models. (a). ANN (b). SVR (c). RF (d) GBDT

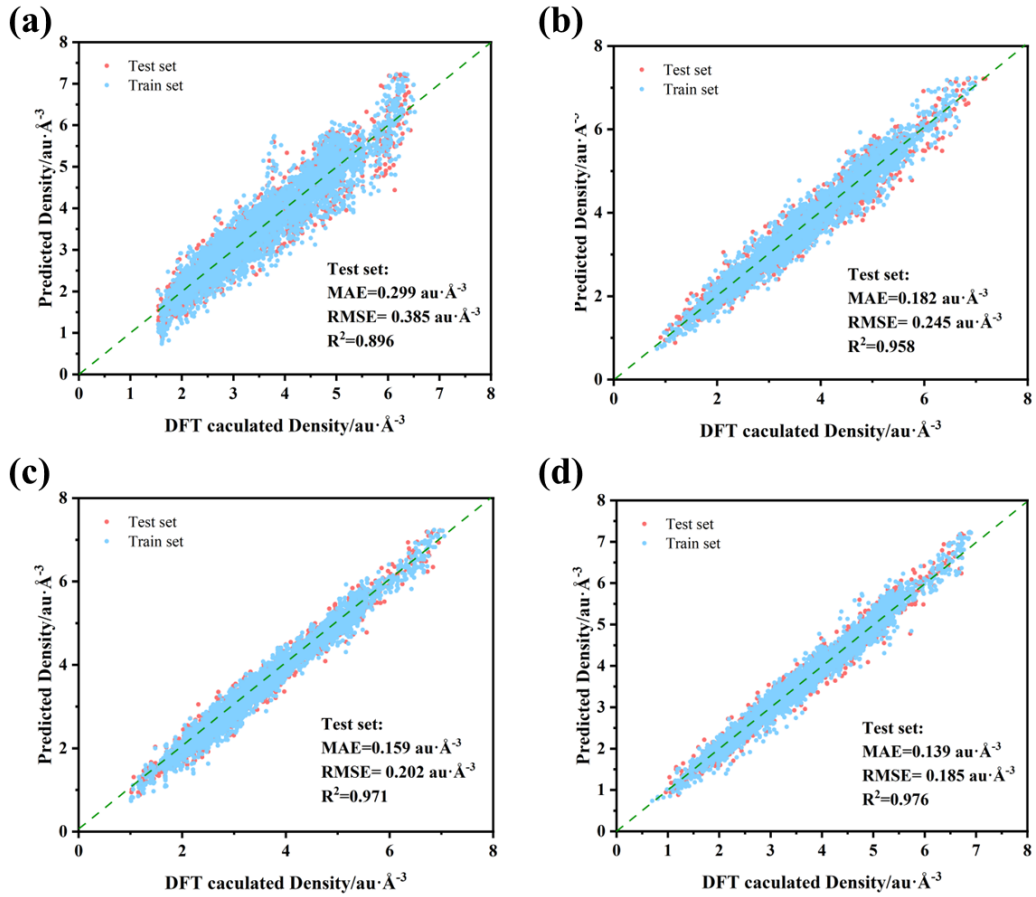

Figure S3. Scatterplot of predictions of perovskite density by classical machine learning models. (a). ANN (b). SVR (c). RF (d) GBDT

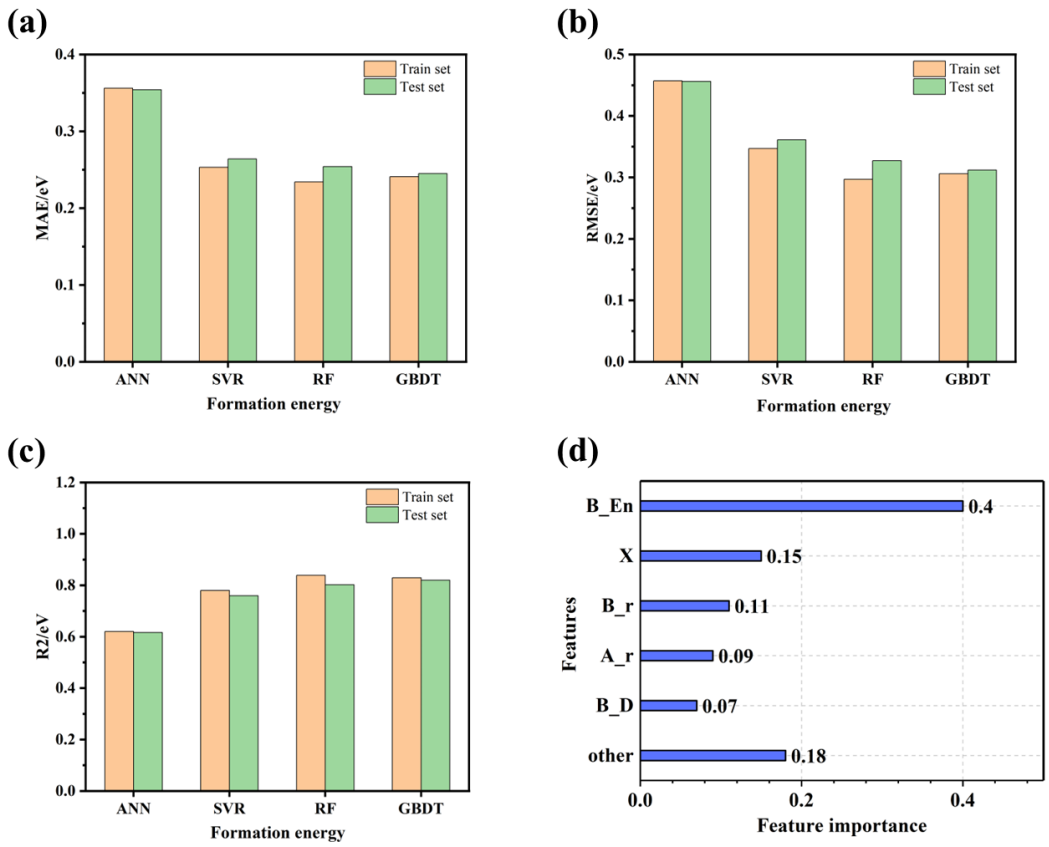

Figure S4. Train and test set metrics and feature importance for prediction formation energy of classical machine learning models. (a). MAE (b). RMSE (c). R^2 (d). Feature importance, r for atomic radius, N for atomic number, En for electronegativity, D for dipolar polarizability, and "other" to represent the sum of any additional properties.

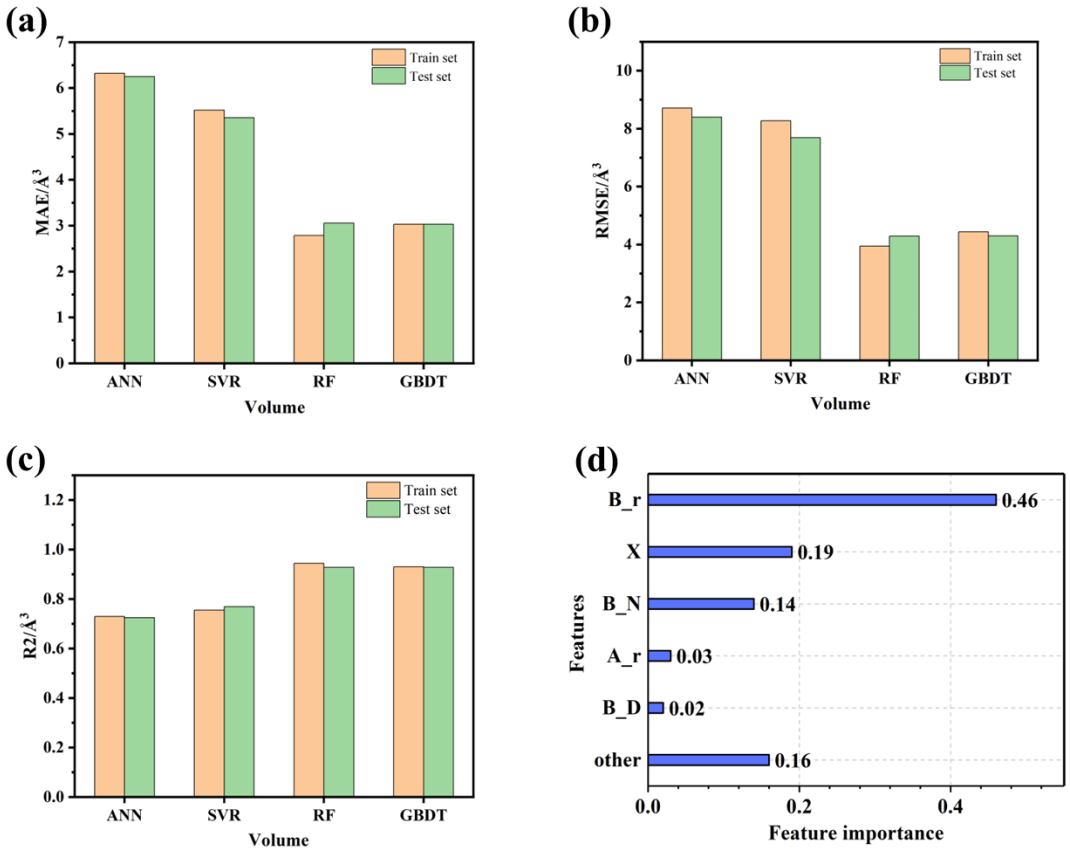

Figure S5. Train and test set metrics and feature importance for prediction volume of classical machine learning models. (a). MAE (b). RMSE (c). R^2 (d). Feature importance, r for atomic radius, N for atomic number, En for electronegativity, D for dipolar polarizability, and "other" to represent the sum of any additional properties.

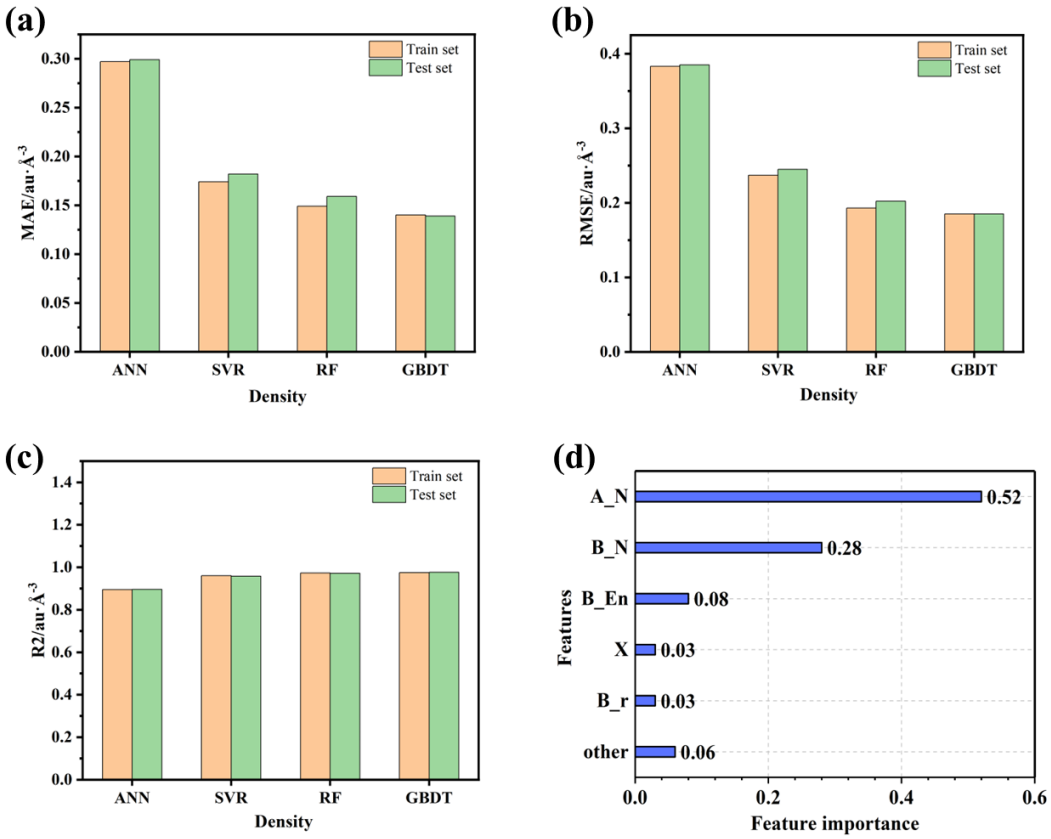

Figure S6. Train and test set metrics and feature importance for prediction density of classical machine learning models. (a). MAE (b). RMSE (c). R^2 (d). Feature importance, r for atomic radius, N for atomic number, En for electronegativity, D for dipolar polarizability, and "other" to represent the sum of any additional properties.